

Georgia Archives
Concept of Operations for the Digital Archives of Georgia
(DAG)

Version 1.4
April 2007

Revision Table

Date	Filename/version #	Author	Revision Description
1/8/2007	ConOps1.0	Winstead	Original draft
1/11/2007	ConOps 1.1	Winstead	Reorganized sections
1/22/2007	ConOps 1.2	Winstead	Incorporate Comments
2/14/2007	ConOps 1.3	Winstead	Incorporate Comments
4/4/2007	ConOps 1.4	Winstead	Incorporate Comments from RPM

Table of Sections

1 Introduction	6
2 Analysis of Proposed System	9
3 Current Environments	11
4 Proposed System	17
5 Operational Concepts	27
6 Operational Scenarios	51
7 Summary of Impacts	64
8 References	66

Table of Contents

1 Introduction	6
1.1 Purpose	6
1.2 Scope	6
1.3 Identification	6
1.4 Document Overview	6
1.5 System Overview	7
1.5.1 Major Operational Functions	8
2 Analysis of the Proposed System	9
2.1 Risks	9
2.2 Disadvantages and Limitations.....	10
2.3 Alternatives and Tradeoffs.....	10
2.3.1 Avoid the Issue	10
2.3.2 Convert Electronic Records to Paper or Microfilm	10
2.3.3 Require Agencies to Preserve Their Own Electronic Records.....	11
2.3.4 Develop a Digital Archives System	11
3 Current Environments	11
3.1 Background, Objectives, and Scope of Current Situation.....	12
3.1.1 Organizational Structure	12
3.2 Description of Current Processes.....	13
3.2.1 Records Transfer	14
3.2.2 Data Management	14
3.2.3 Storage of Records.....	15
3.2.4 Management.....	15
3.2.5 Access to Records.....	16
3.3 Motivation for Digital Archives	17
4 Proposed System.....	17
4.1 Background, Objectives, and Scope of Digital Archives.....	17
4.2 Proposed System Attributes	20
4.3 Proposed System Capabilities.....	20
4.4 Proposed System Interfaces	22
4.5 Operational Policies and Constraints	22
4.6 Description of Proposed System	22

4.6.1 Information Packages	22
4.6.1.1 Operational Environment and Characteristics	23
4.6.1.2 Capabilities, Functions, and Features of the proposed DAG	23
4.6.2 Metadata	24
4.6.2.1 Operational Environment and Characteristics	24
4.6.2.2 Major System Components and High Level Interconnection	25
4.6.2.3 Capabilities, Functions and Features of the Proposed DAG	26
5 Operational Concepts	27
5.1 High-Level Conceptual Overview	27
5.1.1 Workspace	27
5.1.2 Preservation	27
5.1.3 Access	28
5.2 User Level Operations	28
5.2.1. Public Web Interface	28
5.2.1.1 Role-based Toolsets	29
5.2.2 Staff Intranet Site	29
5.2.2.1 Role-based Toolsets	29
5.2.2.2 Example Toolset	30
5.3 Digital Archives Objects	30
5.3.1 Transferred Content	30
5.3.1.1 Capabilities, Functions and Features of the Proposed DAG	30
5.3.2 Converted Content	31
5.3.3 Harvested Content	31
5.3.3.1 Capabilities, Functions, and Features of the Proposed DAG	31
5.4 System Functionality	33
5.4.1 Ingest	33
5.4.1.1 Capabilities, Functions, and Features of the Proposed DAG	33
5.4.2 Content Processing/System Administration	34
5.4.2.1 Interfaces to Systems and Procedures	35
5.4.2.2 Unique Identifier	37
5.4.2.3 Persistent Name	37
5.4.2.4 Authentication	38
5.4.2.5 Version Control	39
5.4.2.6 Access	40
5.4.2.6.1 Search	40
5.4.2.6.2 Request	42
5.4.2.6.3 Cataloging	43
5.4.2.6.4 Reference Tools	43
5.4.2.6.5 Interface	44
5.4.2.6.6 User Support	45
5.4.2.6.7 Data Mining	45
5.4.2.7 Preservation Services	46
5.4.2.8 Storage Management	47
5.4.2.9 Security	48
5.4.3 Content Delivery	48
5.5 Modes of Operation	49

5.6 User Classes and Other Involved Personnel.....	49
5.6.1 Profiles of User Classes.....	50
5.6.2 User Class Capabilities.....	50
5.6.3 Interaction among User Classes.....	51
5.6.4 Other Involved Personnel	52
5.7 Support Environment.....	52
6 Operational Scenarios	52
6.1 Transferring Agency	53
6.1.1 Scenario -Transfer	53
6.1.1.1 Activities	53
6.2 Content Evaluator.....	54
6.2.1 Scenario – Transfer and Ingest.....	54
6.2.1.1 Activities	55
6.3 Content Processor.....	56
6.3.1 Scenario - Process	56
6.3.1.1 Activities	56
6.4.1 Scenario - Administrator	60
6.4.1.1 Activities	60
6.4.2 Scenario - Operations Manager.....	61
6.4.2.1 Activities	62
6.5 End User	63
6.5.1 Scenario.....	63
6.5.1.1 Activities	63
7 Summary of Impacts	65
7.1 Operational Impacts	65
7.2 Organizational Impacts.....	66
7.3 Impacts during Development.....	67
8 References	67
8.1 Standards and Guidelines	67
8.2 Laws and Regulations	68

List of Figures

Figure 1 Operational Model	8
Figure 2 Reference Model	18
Figure 3 Expanded Categories of Information	19
Figure 4 Content Package Information	19
Figure 5 Example Content Packages	24
Figure 6 Proposed DAG Metadata Model	25
Figure 7 System Administration/Content Processing Component	34
Figure 8 Content Processing	35
Figure 9 User Classes	51

Concept of Operations

1 Introduction

The Georgia Archives was established in 1918 by public law 434 (now codified as O.C.G.A. §45-13-40 et seq.). As the official archives for the state of Georgia, the Georgia Archives is responsible for the effective and efficient management, preservation, and use of public records of state government. To continue to fulfill its mandate, the Georgia Archives must respond effectively to the challenge posed by electronic records – digital content and digital objects¹ as they are termed in this document. To that end, the Archives must develop new policies and procedures for the management and care of digital objects. As part of that process, we are developing a digital archives system at the Archives, to preserve and provide access to permanent digital content. The Concept of Operations document is intended to support the development of this system.

1.1 Purpose

The digital archives system is intended to maintain the accessibility of historical digital objects over long periods of time and to maintain their authenticity and reliability as evidence of government transactions.

1.2 Scope

This document describes the desired characteristics of the Digital Archives of Georgia (DAG) from the user's viewpoint. The sections below identify the proposed system, provide a document overview and describe the approach used to generate the document, and provide a brief overview of the system.

1.3 Identification

The proposed DAG will encompass all the associated equipment, facilities, material, software, hardware, policy and technical documentation, business practices, and personnel required for its operations and support at the Georgia Archives, Office of Secretary of State.

1.4 Document Overview

The DAG Concept of Operations serves as a method to communicate high-level system characteristics of the envisioned system to the user, buyer, developer, and other stakeholders. This includes a high-level depiction of system processes

¹ Digital content within the scope of the system is any permanent digital records identified by retention schedule (or other means) and digital objects, such as a web page or data warehouse determined as having historical value but that may not meet the definition of a record. All such digital content is created by a Georgia government entity or its legal agent. A digital object is a single instant of digital content.

and a set of examples of user-level operations. The ideas expressed within are the result of a thorough analysis of the challenges involved in preserving content as expressed in the OAIS model². The Concept of Operations document will be used to generate system specifications documents and technical requirements documents for the development of the DAG.

This document contains the following sections:

- Section 1, Scope, describes the approach used to develop the DAG concept of operations document.
- Section 2, Analysis of the Proposed System, analyzes the proposed system.
- Section 3, Current Environment, describes the current Georgia Archives processes and systems for managing and preserving records.
- Section 4, Concepts for Proposed System, discusses proposed system concepts.
- Section 5, Operational Concepts, provides a high-level overview of the system.
- Section 6, Operational Scenarios, describes various operational scenarios.
- Section 7, Summary of Impacts, summarizes operational, organizational, and other impacts that could be expected to occur during development.
- Section 8, References, lists the reference documentation that was used as a basis to create this document.

The Concept of Operations document is based on the requirements of IEEE 1362-1998, IEEE Guide for Information Technology – System Definition - Concept of Operations (Con Ops) Document. However, the Georgia Archives has reorganized the presentation of the information to the reader.

1.5 System Overview

The DAG will be composed of the necessary technology and business practices that will enable the Georgia Archives to aid selection, ingest, manage, preserve, and provide access to permanent digital records of state government. The DAG will become the primary repository for the permanent digital content of Georgia government.

The proposed system is envisioned as a comprehensive, systematic, and dynamic means for preserving content independent of specific hardware and/or software. When it becomes operational, the DAG will enable Georgia Archives

² The Open Archival Information System (OAIS) model was developed by the Consultative Committee on Space Data Systems and serves as a model for the development of digital archives systems. Refer to section 4.1 for a detailed overview and explanation.

customers³ to access and retrieve the permanent digital records of state government.

1.5.1 Major Operational Functions

Included below is a figure representing the five major operational functions of the DAG.

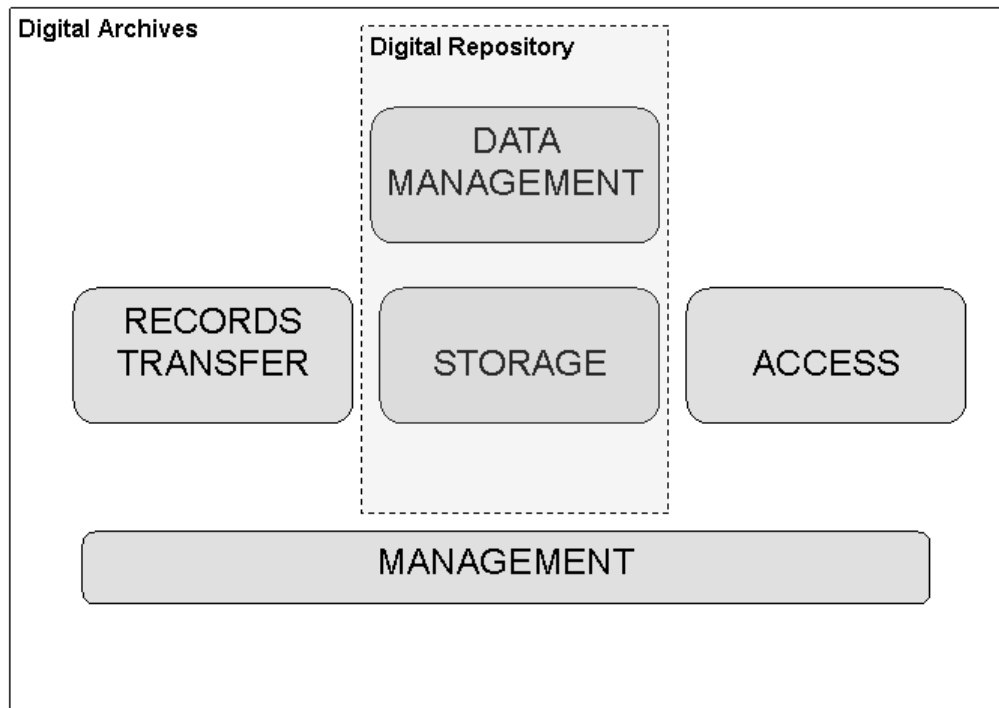


Figure 1. Operational Model

Digital Archives

The Digital Archives includes all five operational functions. Digital Archives refers to the system as a whole, including technical infrastructure and processes for managing and providing access to permanent digital content and its associated metadata.

Digital Repository

Digital Repository includes the Storage and Data Management functions. Digital Repository refers to the actual server where digital content is stored and where the metadata associated with that content is captured and maintained.

³ The customers of the Georgia Archives consist of members of the public and government conducting research into the historical records of the state of Georgia for genealogical, academic, business, or other purposes.

Operational Functions

The Digital Archives specifies five major operational functions:

- *Records Transfer*: The process of ingesting, critically examining, and accessioning (or rejecting) digital content transferred by government agencies.
- *Data Management*: The process of identifying and describing digital content, including the creation of new metadata, as well as maintenance and management of existing metadata.
- *Storage*: The process of preserving content over extended periods of time, including media management, and security.
- *Management*: The process of managing the overall operation of the archive system, including the management of the system hardware and software.
- *Access*: The process of managing the retrieval and viewing of content, including browsing, searching, ordering, and viewing content and its descriptive metadata.

2 Analysis of the Proposed System

Various improvements, disadvantages and limitations, and alternatives and trade-offs considered are covered in this section.

2.1 Risks

Electronic records can be read only if the proper hardware and software exists. The pace with which technology changes make it unlikely that today's electronic records will be readable in twenty or thirty years, let alone 200 years. Just 25 years ago, eight-inch floppy disks were still a common method of storing electronic files. Today, no computer in state government is capable of processing such a disk; and such a computer, if one were found, would no longer contain the software needed to render the files⁴. Yet state agencies create records every day that must be available to citizens and historians for decades—even centuries—into the future. Many records created in electronic form have long term legal, administrative, and historical value: Governor's correspondence files, e-mails, press releases, and Executive Orders; laws passed by the General Assembly (which shortly will be created and processed in electronic form); Department of Transportation maps and plans; even the photographs used by Tourism agencies

⁴ Even if an eight-inch disk could be found, it is unlikely that a modern operating system would recognize the hardware.

to document and advertise the state—photos which have always provide a historical record of changing Georgia—are now created in electronic form and may well disappear within decades.

2.2 Disadvantages and Limitations

Disadvantages and limitations related to developing the proposed DAG include:

- Potentially high development costs;
- High costs associated with security;
- Staff anxiety brought about by changing and new responsibilities; and,
- Impact on transferring agencies (resources required to prepare for transfer of materials to Archives, greater content management responsibilities).

2.3 Alternatives and Tradeoffs

In an analysis of alternatives performed by the Georgia Archives, four alternatives and the subsequent tradeoffs were considered:

- Avoid the issue
- Require conversion of electronic files to paper or microfilm
- Require government agencies to preserve their own electronic records and make them accessible
- Develop a Digital Archives system

2.3.1 Avoid the Issue

This alternative would allow the Georgia Archives to ignore the issue and continue a business as usual attitude. Agencies would continue to retain or destroy their digital content in whatever manner they currently do. Within available staffing levels, assistance and consultation would be available to state and local government agencies on the preservation and retention of electronic records.

While this option is definitely attractive as it requires no additional monies or staff for operations and does not change current business practices, the tradeoff is that it fails to meet the Archives mandate and mission to preserve and provide access to the historical records of the state, and it fails the citizenry of the state by jeopardizing their rights.

2.3.2 Convert Electronic Records to Paper or Microfilm

Under this alternative, state agencies would be required to convert digital content to paper format in order to meet retention and archival requirements for long-term storage. The paper records or microfilm would be eligible for transfer to the

Archives. The advantage here is the demonstrated ability to preserve both media for centuries.

The tradeoff here is that we lose the tremendous search and browse capability of the electronic record. Data mining of the vast amounts of data being collected and retained by agencies also becomes impossible. Databases and other dynamic objects would be rendered virtually useless by a snapshot and print strategy. In addition, the requirements for storage space would quickly outstrip the Archives storage capabilities and agency costs for conversion would quickly make this a less than desirable option.

2.3.3 Require Agencies to Preserve Their Own Electronic Records

This alternative will require state agencies to preserve and provide access to their digital content. Each agency would develop its own solution(s) to the challenge of preserving records, resulting in inconsistent systems providing the citizen with uneven access to public records. Lack of standardization and centralization would also result in a duplication of effort and the expenditure of significantly more funds than should the Archives take on the role on behalf of the state. Additionally, should the Archives fail to meet its mandate in a digital environment, the continued need for an archival program to preserve a finite collection of paper materials comes into question as budgets grow tight.

2.3.4 Develop a Digital Archives System

This alternative proposes the design and development of a Digital Archives system for the preservation and access of government digital content. Agencies would transfer (or allow the harvest of) digital content (documents, emails, publications, databases, web sites, etc) to the system for storage, further description, preservation, and access. Working with the Georgia Technology Authority, the Archives would issue technology standards to promote the creation of archives-ready records that would be more cost effective for the state to preserve. This standardization added to the centralized storage of historical digital content, would eliminate duplication of effort and maximize the funds dedicated to the preservation of such records. In addition, the Archives would continue to meet its mandate of preserving the history of the state and providing access to historical records.

The Georgia Archives has elected to pursue this alternative.

3 Current Environments

This section describes the background, objectives, and scope of the current situation at the Georgia Archives and the systems in use; operational policies or constraints; modes of operation; user classes and other involved personnel; and the support environment.

3.1 Background, Objectives, and Scope of Current Situation

The Georgia Archives was established in 1918 public law 434 (now codified as O.C.G.A. §45-13-40 et seq.). As the official archives for the state of Georgia, the Georgia Archives is responsible for the effective and efficient management, preservation, and use of public records of state government.

Public records serve as documentary evidence of government actions, both past and present, and date back to the establishment of the colony of Georgia in 1732. Public records are created by the agencies of Georgia government including the executive, judicial and legislative branches, corporate authorities, health institutions, and educational institutions.

Currently, the Georgia Archives has acquired very few electronic records. All were brought in on physical media, i.e., CD-Rs, and all are maintained on the original media with no intervention. As the Archives plans for the acquisition and preservation of electronic records, it must develop new policies and procedures for the management and care of digital objects. As part of that process, we have documented our existing processes so that key activities are carried forward into the new system environment.

3.1.1 Organizational Structure

The Georgia Archives is a division of the Office of Secretary of State. The division is managed by a director, deputy director, and four assistant directors. The division is divided into three program units – Public Services, Archival Services, and Records & Information Management Services.

Each program unit manages a different aspect of the Archives mandate:

- **Public Services.** The Public Services unit is responsible for promoting the Archives and providing access to the collections. Specifically, this unit includes:
 - Public Reference Room including the Microfilm Library and the Original Documents Reading Area (ODRA),
 - Volunteer and intern programs,
 - Friends of the Archives, and
 - Educational programs.
- **Archival Services.** The Archival Services unit is responsible for the physical and intellectual management of the collections in the building. This includes:

- Preservation and display activities
 - Reformatting labs (both microfilm and imaging)
 - Inventory control and description
- Records & Information Management Services. The Records & Information Management Services unit works with Georgia government agencies (both state and local) to ensure best-practice records management in government. Specifically, this unit is involved in:
 - Operating a records center for the storage of temporary government records and media security vault for the storage of disaster backups of vital and historical records (for state and local government agencies)
 - Providing training and publications in records management
 - Appraising records and promulgating retention schedules
 - Leading the digital archives project

The Archives operates two facilities: The Georgia Archives building at 5800 Jonesboro Road, Morrow, Georgia, and the State Records Center building at 1050 Murphy Avenue SE, Atlanta, Georgia. The Archives has a total of 45 staff positions with the majority (39) being based at the Morrow location.

3.2 Description of Current Processes

This overview is limited in scope to only those activities relating to the acquisition, care, description, and access of the holdings of the Archives. In other words, to the functions that will also be present in the Digital Archives. These functions, with the exception of Access, are largely focused within the Archival Services program unit.

- *Records Transfer*: The process of transferring, examining, and accessioning paper records or other media transferred by government agencies for permanent retention.
- *Data Management*: The process of identifying and describing records in the holdings of the Archives in order to create and maintain finding aids and other access tools.
- *Storage*: The process of preserving records over extended periods of time, in environmentally controlled storage areas, vaults.
- *Management*: The process of managing the overall operation of the vaults, including monitoring of environmental and storage container conditions.

- **Access:** The process of managing public access of records, including the facilitating of finding aids, retrieval of records through ODRA, monitoring the viewing of records in ODRA, presentation of digital content on the web site, and reproduction of copies upon request.

The objects managed by the Georgia Archives include: paper records, bound volumes, microforms (16mm, 35mm, aperture cards, and microfiche), CDs, audio cassettes, video tapes, photographs, and motion picture film. These disparate items are controlled through the use of both manual (paper) finding aids and electronic tools (inventory management database).

3.2.1 Records Transfer

Legal and physical custody of permanent records is transferred to the Archives through the use of a transmittal form. The transmittal form is completed by the agency and submitted to the Archival Services unit as notification that eligible records are ready for transfer to the Archives.

The transmittal form can be submitted electronically by the agency as an Excel spreadsheet or a MS Word template. Information gathered on the form includes:

- Agency, division and program unit title
- Address
- Custodian of records
- Staff member coordinating transfer (usually the Records Management Officer)
- Series title of records
- Date span
- Retention schedule identifying records as permanent
- Shipment list (usually at the container level but can be at folder level)
- Total number of items (containers) in shipment
- Type of container (cubic foot box, volume, etc)
- Accession number (left blank until the records are accepted by the Archives)
- Signature of Archives staff member accepting transfer

Archives staff verifies the correctness of the information on the transmittal form and schedules a pick-up or delivery date for the records. As the records enter the Archives, they are inspected to verify the contents of the shipment, to determine the condition of the records, and to identify processing requirements (re-foldering, re-boxing, and separation of media). The containers are then moved into a predetermined work area – preservation labs, reformatting labs, processing rooms.

3.2.2 Data Management

Once records are acquired through the records transfer process, the Archives must gather information about the records in order to tell others what records we have, locate the records, and provide access to them. This is done during archival processing.

Processing involves the physical and intellectual work flows that enable the Archives to index and classify records in the Archives descriptive control system. For preservation purposes, records may be re-folded, re-boxed or repaired during processing. At the same time staff are gathering information on the filing system, subjects included within the records, and key identifiers for the records. Finally, a location number identifying where the container will reside in the Archives vaults is assigned to the record. This information is added to the information initially acquired on the transmittal form – agency, division, unit names; series title; date span; retention schedule; accession number (control number that identifies and links each shipment); and contents list. All of this information is entered into the Archives descriptive and inventory control system – Context by Tower software. As additional records are transferred, the descriptive and inventory control system is updated.

3.2.3 Storage of Records

The Archives currently manages and maintains over 76,000 cubic feet of records in a variety of media (microfilm, audio/video tapes, and paper). This total equates to over 100,000 individual objects. The descriptive and inventory control system of the Archives links the physical location of an object with the indexing terms required for the general public to access the record. (The Archives does not have public access to TRIM, the public uses Online Descriptive Inventories and paper versions to request government records and staff at the reference desk use TRIM to locate the records). Within the Archives building, there are four vault floors, each with controlled access. Records are stored by container type to maximize the number objects that each floor is able to store. The container type, unique identifying number and location number (identifies vault floor, row, and shelf) are used in combination to distinguish the thousands of objects within the Archives. Vault spaces within the Archives are maintained at optimum temperature and humidity levels to ensure the preservation of the media upon which the records are stored.

3.2.4 Management

Long-term management of the records involves maintaining and monitoring the physical infrastructure around the records. Temperature and humidity are documented and tracked for fluctuations that would over time result in the deterioration of collections. As needed, containers are replaced. In addition, reformatting of the certain collections may be considered as more and more of the Archives holdings are made available electronically.

3.2.5 Access to Records

The Archives fulfills records requests and provides copying services to both the general public and to Georgia government.

The request and copying services are provided to the general public via the Internet and onsite:

- The Archives is open to the public 5 days a week (Tuesday – Saturday) from 8:30am to 5:00pm. Requests for information may also be submitted via the Archives web site through an email reference request application called Ask an Archivist and by phone. Treasures of the Archives (i.e., Royal Charter for the Colony of Georgia) as well as popular collections (i.e., Confederate Pension Applications) are available via searchable databases as digital images.
- Members of the public visiting the Archives are asked to register at the customer service and show a photo id. They are then directed to a security desk where notebooks and other items being carried into the reference areas are searched for contraband (knives, ink pens, food and other items). Lockers are provided to secure items such as purses and briefcases.
- Members of the public have open access to the library stacks and to the reference microfilm collections. If original records are requested, a records request form identifying the record group, series, and box numbers is completed and submitted to the Original Documents Reading Area (ODRA) where records are viewed.
- Ordered records are delivered to ODRA within 30 minutes. Five records requests may be submitted at once but only one unit may be used at a time. (the five records requests is up to five boxes from the same series and/or up to five different series with one box or any combination thereof)
- Paper photocopies, microfilm printouts, or digitized scans or prints may be ordered of original records. Photocopy requests are usually completed the same day but the Archives reserves the right to fulfill larger orders within a period of ten working days.

Similar services are provided to government agency staff:

- Government agencies may request records by phone, fax, email or in person. Typically, the agency creator is requesting access to permanent records housed at the Archives. Copies of the records may be requested. Up to fifty copies are free.

- Although, closed to the public on Mondays, government agencies requiring access to records may arrange to be provided access. Agency records are not returned to the agency after they have been transferred to the Archives but agency staff may access the originals in the ODRA and may request copies.

3.3 Motivation for Digital Archives

Implementing a Digital Archives system will require business process re-engineering of all current work flow within the Archives. The system, when implemented, will completely change the way the Georgia Archives operates, providing new functionality and capabilities for the preservation and access of all historical content – but particularly of digital content. The full extent of the capabilities will not be known until the completion of the systems analysis and design phase of the program and will be addressed at that time. It is anticipated, however, that the proposed system will offer numerous benefits to the Georgia Archives and its end users. These benefits may include:

- Consolidated content management lifecycle administration and streamlined internal workflow;
- More involvement with agencies in the early phases of the content creation and management lifecycle;
- New tools to support processing and evaluation of content:
 - Tools to aid in the ingest process,
 - Tools for preservation and access, and
 - Tools for the creation of descriptions and metadata;
- A wider variety of available content within the holdings of the Archives;
- Faster access to content;
- Enhanced capabilities for searching content;
- The ability to service additional end users;
- Increased responsiveness to end users;
- Remote access to content; and,
- The harvesting, locating and preservation of content that would otherwise be lost.

4 Proposed System

4.1 Background, Objectives, and Scope of Digital Archives

The basic terms and concepts used in the DA model are largely drawn from the Consultative Committee for Space Data Systems' Reference Model for an Open Archival Information System (OAIS). The OAIS model was developed by the Consultative Committee on Space Data Systems (CCSDS) with broad input from other communities. It was issued by the International Standards Organization (ISO) in 2003 as standard ISO 14721:2003: *Space data and information transfer systems – Open archival information system – Reference model*. OAIS is a

domain neutral reference model with characteristics broadly applicable to the management of any information over time. The OAIS model has been adapted and used in other research collaborations and provides the scalability, extensibility, and interoperability required for a system of this magnitude. This model does not prescribe an implementation. Using the OAIS as a reference model begins the process of defining what is necessary to achieve the strategic objectives of the Georgia Archives for establishing the content management lifecycle processes for the management and preservation of content of all types. It is recommended that the system be an integrated system that provide OAIS foundation services (see figure 2) such as ingestion of content, storage of that content in the form of electronic records for as long as needed, content management, and the ability to provide access to the content from anywhere on demand.

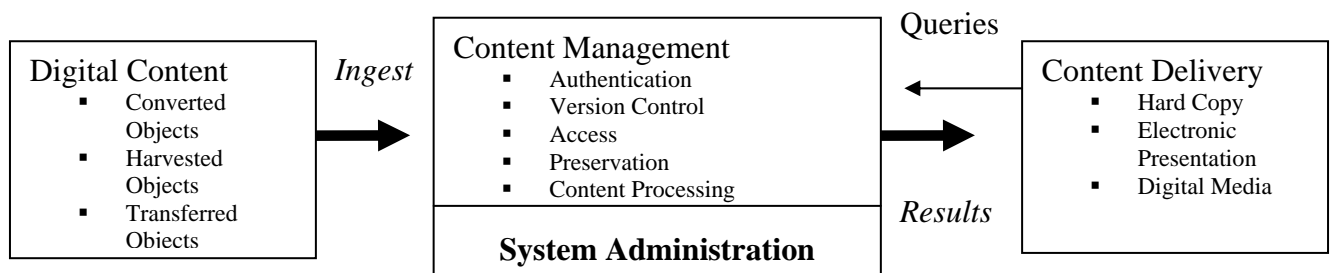


Figure 2. Reference Model

From a design concept, the architecture of the system must

- Eliminate conceivable “single” points of failure;
- Allow media, software and hardware to be interchangeable over time as they fail or become obsolete;
- Support diversity among its components to:
 - avoid monoculture vulnerabilities,
 - allow for incremental replacement,
 - avoid vendor lock-in; and
- Detect errors and failures even when the information is not accessed.

To meet user expectations, the architecture for the DAG should be able to:

- Ingest digital content from state agencies.
- Manage content via an enterprise class content management system.
- Support an effective content delivery capability.

The information managed by the system has two characteristics divided into categories: content information package (CIP) and business process information (BPI). Business process information is administrative, non-content specific information that is used within the business process and package description (PD) to support finding aids and data mining. CIP directly relates to the content and is ultimately used in the dissemination and preservation of the content to the end users. CIP includes the actual object, descriptive metadata, administrative

information (such as creator and security policies), and preservation metadata added by the digital archives to facilitate eventual migration of the object.

These information categories are further defined in a manner that is consistent with the OAIS and the definitions developed for the DAG. (shown in Figure 3)

BPI	Package Description (PD)		Administrative Information (AI)	
	Information about the content package information (CPI) that is used by finding aids includes select external information.		Customer profiles, statistics, security policy, etc.	
CIP	Digital Archives Object	Representation Information (RI)	Packaging Information (PI)	Preservation Description Information (PDI)
	An object composed of a set of bit sequences	Metadata that maps digital archives object into more meaningful concepts	Information used to bind and identify components of the AI	Metadata for adequate preservation of the content package

Figure 3. Expanded Categories of Information

The RI will include original metadata and any other information, objects, or applications that are required to render the content at a specific level of accuracy. The RI should also include enough metadata for certification, version control, access, and preservation to take place. Without adequate information within the RI, decisions on the metadata content of the PDI cannot be made. For example, a harvested content package (a web site) may contain a web page (the object); information about the web page, the software and version, and how it is structured and formatted (RI); and metadata that contains information about the agency creator, and publication date (AI and PI). (Figure 4 illustrates content package information)

Digital Archives Object	Representation Information (RI)	Packaging Information (PI)	Preservation Description Information (PDI)
An object composed of a set of bit sequences	Metadata that maps digital archives object into more meaningful concepts	Information used to bind and identify components of the AIP	Metadata for adequate preservation of the content package

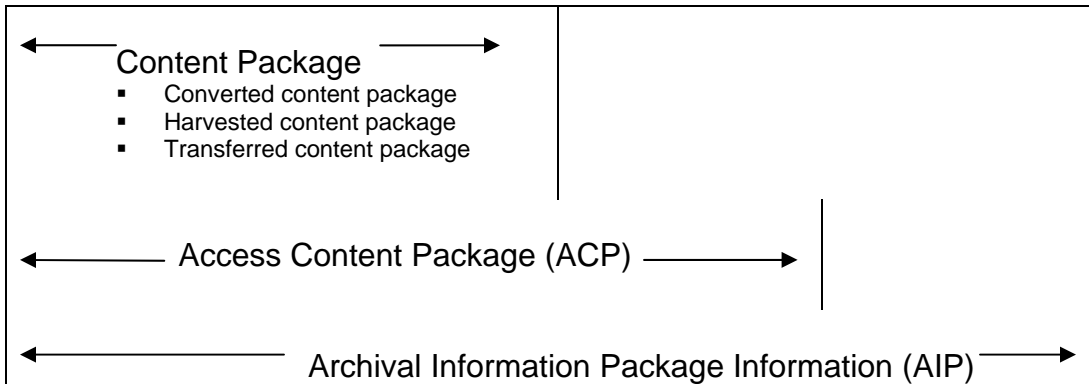


Figure 4. Content Package Information

Scope definitions:

Content within scope of the DAG –

Content within the scope of the system is any permanent digital records (identified by retention schedule or other means) created by a Georgia government entity or its legal agent.

4.2 Proposed System Attributes

From an overall system perspective, the DAG should possess the following attributes.

- Infrastructure independence: an architecture that allows preservation of content independent of any specific hardware and software that was used to produce them;
- Modularity: ability to use plug-in components that can be replaced with minimal impact to remaining components as workload and technology change;
- Scalability: capable of accommodating growth and managing differing sizes of repositories and ever increasing volumes of content;
- Extensibility: be able to handle additional kinds of content over time, not limited to specific types of records; and,
- Flexibility: enable the Georgia Archives to tailor content-based services to suit its customers' needs and enable the Georgia Archives to implement progressive improvements in its business processes over time.

4.3 Proposed System Capabilities

To meet strategic objectives, the Georgia Archives must integrate its solution for preservation and long-term access to content with lifecycle management of that content throughout state government. To meet the challenges of today and the future, the DAG should be able to:

- Accept the transfer of digital records (content) in a wide variety of formats as they were created or stored by the agency and the flexibility to easily adapt to future file formats;
- Ingest, preserve, and provide access to that content;
- Store content in a manner that is independent of any particular hardware and software component over long periods of time;
- Scale in order to store and preserve content based on the predicted digitizing of existing paper archival collections and the harvest of state government web site content;
- Provide access to the content in electronic form for all users based on established user rights and privileges, thus ensuring that the system users are able to access all of the content that they are entitled to see;
- Provide access to the content in a manner consistent with current technology and the changing expectations of diverse user communities;
- Adapt to changing technology in order to continue to provide access to and delivery of content desired by user communities; and,
- Identify the essential characteristics of the content that is being preserved for the purposes of authentication and certification.

The proposed DAG should provide the following capabilities in support of Georgia Archives content management lifecycle processes.

- Provide end-to-end automated work processes that streamline the content management lifecycle processes for all content;
- Manage the authentication, transfer and ingest of all content;
- Support the end-to-end tracking of all content during the process of transfer and ingest;
- Ensure that content transferred to the digital archives remains free from corruption;
- Accept transfers of content, check that the content conforms to terms and conditions of the transfer documentation, and store them in the system;
- Provide an automated tool to inform the Georgia Archives of any content eligible for transfer that has been submitted by an agency through the web portal;
- Support preservation services so that the original bit stream can be preserved over time;
- Support the description of content so that it is clearly identified, discoverable, and retrievable;
- Enforce restrictions on access and release of content;
- Output copies of content as specified by users;
- Output authentic certified copies of content;
- Monitor system performance;
- Maintain system security; and

- Provide audit trails of system activity.

4.4 Proposed System Interfaces

The proposed DAG will be government-wide and will operate within the context of the Enterprise Architecture. The system will be capable of interfacing with other applications through state government for transfer of content to the Archives, for retrieval of content by its creator, and for content management lifecycle processes in which the Archives interacts with other entities in all three branches of government. The volume and diversity of content that is ingested and disseminated, and the potential for expected heavy use of the system, will have considerable impact on the Archives computing environment.

Interfaces to other Archives systems as well as to other government agency systems will be accommodated by the system. Specific interfaces are yet to be determined and will be described in future technical documents when identified.

4.5 Operational Policies and Constraints

The DAG is required to be policy neutral so that it can support not only the Archives' current operational policies but also changes that are expected to emerge. The DAG will utilize current technology standards and as these standards change over time, objects within DAG will be migrated and converted as necessary.

Constraints as currently known that may impact system architecture or major system components are:

- The system design and implementation is flexible and adaptable to changes in hardware, software, communication technology, processes, policy, personnel, locations, etc.;
- The system is responsive to policy, but is not policy-constrained;
- The system design to the extent possible, is Commercial Off The Shelf;
- The system implementation is within the specified budget and timeframe;
- The system will interface with back-end infrastructure systems at the Georgia Archives (e.g. Tower); and
- The system will adhere to all applicable levels of privacy, security and accessibility policies (division, agency, and enterprise levels).

4.6 Description of Proposed System

This section provides information describing the architecture and functionality of the proposed DAG.

4.6.1 Information Packages

Information packages stored and managed in the proposed DAG are digital objects (content) and representation information, to which descriptive, technical, and administrative metadata are added to enable and facilitate lifecycle processes such as preservation, search, retrieval, and dissemination.

Four types of information packages are created and maintained by the system:

- Submission Information Package (SIP) consisting of the archives-ready content transferred by an agency and ready for ingest;
- Access Content Package (ACP) consisting of the object and metadata to facilitate access functions;
- Archival Information Package (AIP) consisting of a copy of the object and metadata to facilitate preservation processes;
- Dissemination Information Package (DIP) consisting of a copy of the object and metadata to facilitate delivery of the information to users.

4.6.1.1 Operational Environment and Characteristics

Information packages in the proposed DAG will provide the basis for content processing functions including ingest, version control, validation, preservation, and access, as well as for content delivery functions. In order for the system to effectively preserve information and provide permanent access, it must ensure that official content created by government agencies and essential information about that content is managed in an environment that facilitates processing of and access to fully functional information, as well as an array of preservation processes.

4.6.1.2 Capabilities, Functions, and Features of the proposed DAG

The system will recognize related content objects and metadata as information packages. In order for the system to preserve content and assure access, information packages are created to be responsive to the particular demands of system functions. The Submission Information Package, the Access Content Package, and the Archival Information Package are created in ingest. The Dissemination Information Package is created in content processing. In all cases, the packages must be composed of the fullest possible representation of the object, free of the barriers of proprietary formats and associated with metadata appropriate to the function. The information packages used in the system will be iterated across the three sources of content – transferred, harvested, and converted (see figure below for example).

Content Packages	Submission Information Package	Access Content Package	Archival Information Package	Dissemination Information Package
Transferred Object	Digital object as submitted by agency	Validated digital object	Faithful copy of the validated object	Faithful copy of the object rendered for dissemination
Transferred Metadata	Representation information, stylesheets, DTDs, and other submission level elements	SIP metadata + descriptive metadata, packaging metadata for access, content transformation, content management, etc	SIP metadata + PDI for harvested content	ACP metadata + additional descriptive and packaging metadata for dissemination
Converted Object	Single page, 1:1, uncompressed TIFF, no cropping, deskewing, interpolation, or other modification	Files created from converted object and/or other derivatives intended for access (eg, text, JPEG)	Single page, 1:1, uncompressed TIFF, no cropping, deskewing, interpolation, or other modification	Object retrieved from Access Content Storage
Converted Metadata	Full technical information from the conversion following NISO Z39.87-2002; required submission level	Submission level +metadata for access, content transformation, content management, derivation, etc	SIP + PDI for converted content	ACP metadata + packaging metadata for dissemination
Harvested Object	Digital Object as harvested	Digital object as harvested and file(s) as converted for access	Digital object as harvested and file(s) rendered for optimal preservation	Faithful copy of file(s) in the ACP, rendered for dissemination
Harvested Metadata	Representation information and documentation of harvest & transformation(s); submission level	SIP metadata + descriptive metadata, packaging metadata for access, content transformation, content management, derivation, etc	ACP metadata + PDI for harvested content	ACP metadata + additional descriptive packaging metadata for dissemination

Figure 5. Example content packages

4.6.2 Metadata

Actions or processes in the proposed DAG require and/or create information about content in order to be carried out. This information is recorded, stored, and subsequently used as metadata. Metadata is a structured representation of information that facilitates interpretation, management, and location by describing essential attributes and significant properties of information. Fundamentally, metadata describes how, when, and by whom a particular set of data was collected, what the data is, where it resides, and how it is formatted.

4.6.2.1 Operational Environment and Characteristics

Metadata creates a systematic approach to expressing information derived or discerned from the content itself or from processes associated with the content. It encompasses static properties (ie, those related to the specific instance or

version of the content being processed, queried, or preserved) as well as the aspects of the lifecycle of the object, a continuum extending from creation, through system ingest, preservation, content management, access, and use. Metadata is generally classified in several broad categories, according to its function:

- Descriptive – such as bibliographic information describing, classifying, and characterizing the identity and context of the content.
- Technical – describing file format, computer environment, functionality, etc, in which the content was created or acquired and the attributes of the technical environment necessary to render the content meaningfully.
- Structural – describing interrelationships and hierarchies of files and content.
- Administrative – describing rights, ownership, conditions of use, business rules, etc.
- Preservation – information necessary to maintain viability (the bit stream is intact and readable), renderability (translation of the bit stream into a form useable by humans), and understandability (the rendered content can be interpreted and understood by the intended user). Preservation metadata draws heavily on the other four categories.

Metadata in the proposed DAG must record essential properties and attributes which can be mapped to the major elements in the proposed DAG metadata model, which is broadly adapted from the OAIS metadata model.

4.6.2.2 Major System Components and High Level Interconnection

The below figure illustrates the major system components and high level interconnection.

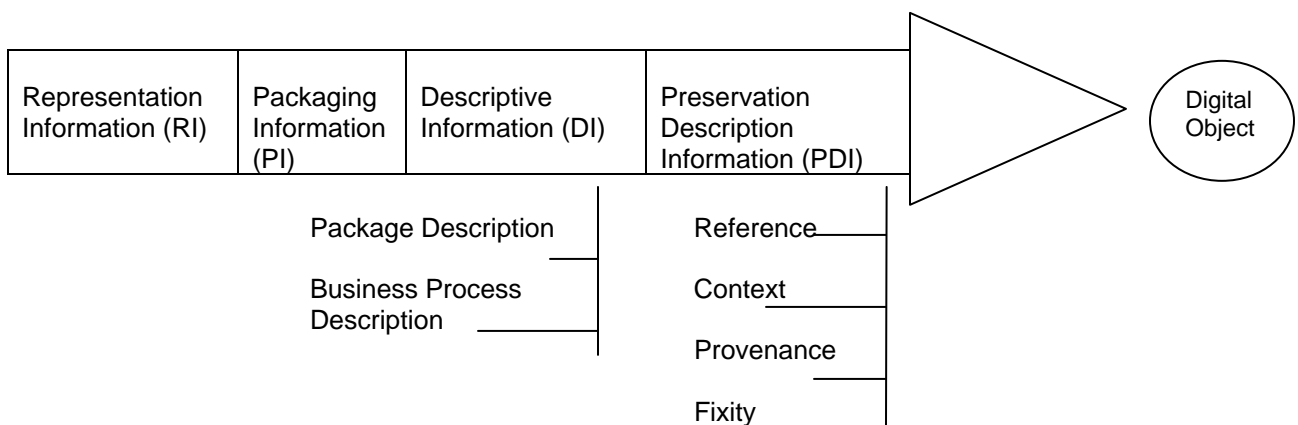


Figure 6 Proposed DAG Metadata Model

Metadata describes, characterizes or supports processes related to digital objects that comprise content within the system. Metadata will also describe and

support business processes within the system that are not related directly to specific content (e.g. system operation and traffic, user profiles, administrative data, or ordering information).

Throughout the lifecycle of digital objects within the system, metadata is captured, collected and recorded either in support of processes or to document their outcome. These actions relate to the OAIS model, in which an information package is formed by adding successive layers of metadata to the content information.

The metadata required for these information packages can be categorized into the four major areas above:

- RI allows content to be rendered and understood. Without it, the digital object is merely an undifferentiated bit stream. Although it is not necessarily expressed as metadata for use in rendering, it must be in order for information to be passed to other functions in the system.
- PDI is necessary to maintain viability (the bit stream is intact and readable), renderability (translation of the bit stream into a form useable by humans or machines), and understandability (the rendered content can be interpreted and understood by the intended user). It is composed of descriptive and technical elements which record identity, fixity, provenance, context, technical environment, lifecycle events, and preservation actions.
- PI is used to bind and identify the components of the information package into a unit.
- DI facilitates access to and management of content. It is composed of elements which record identity, responsibility, version, fixity, structure and technical attributes, administrative data, reference information, and lifecycle events.

4.6.2.3 Capabilities, Functions and Features of the Proposed DAG

The system must effectively recognize and interact with a wide variety of established practices and standards for metadata in multiple schemas that are in common use in Georgia government. Implementation of the system must define methodology for selecting and integrating appropriate schema based on requirements of use and users.

The system should effectively recognize and interact with established practices and standards for metadata in multiple schemas and provide a capability for mapping between schemas for differing functions.

5 Operational Concepts

This section provides information that explains the operational concept of the proposed DAG. This includes a high-level depiction of the processes, a set of examples of user level operations such as toolsets, system support, and test and training activities.

5.1 High-Level Conceptual Overview

The DAG will include various operational functions and processes for managing and accessing permanent digital objects and their associated metadata configured into three distinct storage components:

- Workspace component for temporary housing of objects during the records transfer process;
- Preservation component, aka the Digital Repository, for the long-term permanent storage and management of digital objects; and,
- Access component for temporary housing of records during use.

Each component of the DAG is physically separate from the other, secured via firewalls and access restrictions. Underlying these three storage components are the administrative functions for managing the DAG.

5.1.1 Workspace

Workspace storage encompasses the processes of Records Transfer – receiving, assuring quality, processing, importing, and certifying custody of digital objects. Digital objects will be transferred to the DAG using standard file transfer technologies, such as File Transfer Protocol (FTP) or web services. The objects will be scanned for viruses and validated. A unique identifier consisting of a record id and an accession (shipment) number will be assigned at this point. Administrative and descriptive metadata will be added to and modified as the object moves through the various processes within each of the components of the DAG.

5.1.2 Preservation

Preservation storage encompasses the processes of creating and managing metadata, retrieving data, maintaining data integrity, allowing for operations, storing digital objects, and managing storage media. Three versions of an object will be stored and monitored within the Preservation component, also known as the Digital Repository. The original object transferred by the agency will be encapsulated in XML⁵ and preserved along with a preservation copy stored as a non-proprietary file format. For example, if an agency transfers JPEG or PDF images to the Digital

⁵ One key challenge in utilizing XML as a preservation tool is the proliferation of XML standards for specific professions or needs – Justice XML, Legal XML, Student Information Systems XML, etc.

Archives System, the original JPEG or PDF will be preserved and a TIFF version created as the preservation copy. A third version for publication to the Digital Archives web page will also be created. This version may be redacted to conceal confidential information.

5.1.3 Access

Access storage encompasses the processes needed to provide access to information – browsing and searching, indexing, responding to queries, displaying objects, ordering objects, registering users, and managing access. The public web interface is the primary architectural component of this function. Members of the public will interface with the DAG solely through the Access component. Only logged web addresses will be allowed to query the storage array for the retrieval of descriptive metadata and presentation copies. Presentation copies may be redacted to conceal confidential information.

5.2 User Level Operations

The Digital Archives system will employ a web interface for agency and citizen access. Staff working with the DAG will utilize a customizable intranet interface with a series of toolsets that enable the staff member to complete various tasks such as virus scanning, ingest, metadata encapsulation, and searching.

5.2.1. Public Web Interface

The 'homepage' for the DAG is the general public user's webpage. Members of the public will interface with the DAG solely through the Access component. Only logged web addresses (computers logged into the web through a service provider – i.e, AOL – this prevents spidering of the site) will be allowed to query the storage array for the retrieval of descriptive metadata and presentation copies. Presentation copies may be redacted to conceal confidential information. Copies may be printed from the web page by the patron. Should certified copies be required, the patron will be pushed into an e-commerce or shopping cart appliance in order to facilitate the order. Cross-references to other holdings (paper, microfilm, etc.) of the Georgia Archives will link the patron into directions to the facility and information on mail requests for reference.

Agencies will also access records via this same portal. Only devices within the state domain will be allowed access to un-redacted information. Access will not be provided to authorized staff logging in from a computer not within the state domain, even if login and password are authentic⁶. Authorized agency staff will be assigned a login and password in order to access the DAG. The identity of that agency staff member would be authenticated prior to allowing access into the system.

⁶ Provisions will be made to allow staff working from home to access the system if they are connected to a state network via VPN.

5.2.1.1 Role-based Toolsets

State agency staff will interface with the DAG through an intranet portal containing a series of customizable toolsets that enable the staff member to transfer content to the system and search records within the system. Transfer toolsets will prompt the agency staff member to provide necessary provenance and descriptive information about the content to be transferred. Only authorized users with transfer authority will be provided access to this toolset.

Search toolsets will prompt agency staff to provide necessary search criteria for searching and/or browsing agency records. Logons will be established to ensure that confidential information is provided to only those staff members with appropriate levels of access. Authorized agency staff will also have access to 'related' information from other agencies through access to a data warehouse. The warehouse would access data from multiple agencies and provide reporting tools for manipulating the data. Upon logging in to the agency portal, the individual would select search criteria that would determine what would be searched -- agency records and data only or a mixture of related information.

5.2.2 Staff Intranet Site

Archives staff will interface with the DAG through an intranet portal containing a series of customizable toolsets that enable the staff member to ingest, describe, and search the system. After logon, staff is presented with their role-based toolsets. The toolsets include functionality and notices that are relevant to the staff member's responsibilities and prevent access to functionality in which they are not authorized.

5.2.2.1 Role-based Toolsets

The staff portal is structured into a series of toolsets. These toolsets correspond to the processes of ingesting, describing, and accessing content in the DAG. The toolsets are divided into three groups: workspace, preservation, and access.

- Workspace toolsets: includes processes of relating to records transfer – receiving, assuring quality, processing, importing, and certifying custody of digital objects.
- Preservation toolsets: includes the processes of creating and managing metadata, retrieving data, maintaining data integrity, allowing for operations, storing digital objects, and managing storage media.
- Access toolsets: includes the processes needed to provide access to information – browsing and searching, indexing, responding to queries,

displaying objects, ordering objects, registering users, and managing access.

5.2.2.2 Example Toolset

The layout of the toolset includes:

- Generic tabs which provide access to general-purpose tools common to all toolsets.
- Toolset specific tabs which provide access to the many pages that might be included in a staff's toolset. This would include tabs for all functionality to which the staff member is authorized.
- Main body portlets which provide the initial layer of business functionality. These portlets could include links to specific tools; user-specific tasks and notices; and, miniature web forms for request submittal.

To access a specific component of business functionality, the staff member selects the appropriate tab, and then clicks on the action with the portlet.

5.3 Digital Archives Objects

The digital object is the primary unit of the DAG. Digital objects may include transferred content, converted legacy documents, and harvested materials.

5.3.1 Transferred Content

Transferred content, or Submission Information Package (SIP), is defined as content intentionally submitted to the Georgia Archives by a government agency. The SIP for transferred content will include the digital object received from the agency as well as corresponding agency processing information (agency number, authorized individual, reference information, etc) and additional metadata. The Archives must establish best practices, standards, and templates for capturing transferred content, including metadata to capture the agency's information. The system must be able to accept transferred content that is provided in a wide variety of formats and media, with the Archives being able to convert the content (if needed) into a SIP that is compliant with the best practices reference above. Agencies will use an interface to the system that includes a validation process to input requirements into the proposed DAG. The validation process will contain Archives best practices and templates for content structure and required metadata.

5.3.1.1 Capabilities, Functions and Features of the Proposed DAG

- Agencies to directly interface with Georgia Archives
- Capability to utilize Archives-furnished templates, automatic composition tools, and expert services including pre-validation of digital content (confirming that requirements are complete).

- Ingest process must capture all agency processing information (e.g. the system must capture all relevant metadata the agency provides).
- Content packages may be pushed or pulled from Archives servers, agency servers, storage devices, etc.
- Electronically accept, store and transfer content in a wide variety of formats as they were created, with the flexibility to easily adapt to future formats.
- Identify the essential characteristics and structure of the content being processed along with its associated Representation Information (RI).

The proposed system must be able to handle additional types of digital content (e.g. audio, video, etc.) and not be limited to the formats that exist today.

5.3.2 Converted Content

Converted content is digital content created from a tangible product, such as a paper document. While scanning is one method of conversion, there may be other means by which content is converted, such as by manual text encoding. Other formats of non-text based content include analog audio and video. Specifications will be developed on a case-by-case basis for the creation of these files.

Regardless of conversion method, such projects must produce digital facsimiles and necessary metadata for the system.

5.3.3 Harvested Content

Harvested content is content that is gathered from state agency web sites. Discovery, assessment, and harvesting tools will be used to create a SIP. The discovery tools will locate electronic content and provide information to the assessment tool. The assessment tool determines if the content is eligible for transfer and whether versions of the content already exist in the system. If other versions exist, the tool establishes appropriate relationships with the versions. The harvesting tool gathers the content and available metadata.

5.3.3.1 Capabilities, Functions, and Features of the Proposed DAG

Discovery Tools: tools must discover content based on rules and instructions derived from criteria which define a website within scope of the Archives collecting policies. Functionality of these tools includes:

- The flexibility and ability to define and revise discovery rules and instructions
- The ability to recognize all file types that may reside on the website (e.g. PDF, HTML, audio, video, proprietary word processing software, dynamic content, rich media, XML, etc)

- The ability to discover and interpret metadata associated with a given file.
- The ability to discover, identify, and characterize information on agency web sites, including:
 - Deep web information
 - Query-based databases
 - Agency content management systems
 - Dynamically generated web pages
 - Content contained on FTP servers
 - Content contained behind proxy servers
 - Content contained behind firewalls

Assessment Tools: tools must exist to assess discovered content prior to harvest. Staff may assist in determining if the harvested content is in scope. The specific requirements for assessment tools include:

- The ability to filter out information or resources that do not fall within scope
- The ability to discern whether a particular file already exists within the DAG
- If the file exists in the DAG in another format or version, the assessment tools must be able to associate the harvested content with the existing content
- Provide quality control functions to test accuracy/precision of rule application and to cycle results into rule creation/refinement

Harvesting Tools: tools must be able to systematically capture digital objects and RI. Upon capture, the harvested content will be passed to the assessment tools, where it will be evaluated against a set of rules that define an object within scope. The discovery tools report to the harvesting tools the location of eligible sites. This reporting is accomplished through the exchange of harvesting rules. The rules will include the location and structure of the object along with any information that is pertinent to the act of harvesting. The harvester will then use the rules to capture objects and RI and the content will be transferred to the assessment tools, where it will be packaged for ingest.

The following lists capabilities, functions and features the harvesting tool must possess:

- The tool must be able to harvest, store, and add metadata to multiple records simultaneously. The metadata should include when and where the harvested content was captured. It should also include date and time of capture
- The harvesting cycle, or how often harvesting takes place, should be configurable
- The tool will be able to physically capture a copy of the object and RI, including original metadata and any other information, objects, or

applications that are required to render the content at a specific level of accuracy

- The tool will be able to capture all file types
- The tool will have the ability to capture and interpret all metadata associated with a given electronic publication

5.4 System Functionality

The following sections describe the functionality of the system with respect to the major components of the OAIS reference model – ingest, content processing/system administration, and content delivery.

5.4.1 Ingest

Ingest is the function that accepts the Submission Information Package (SIP) and prepares the SIP for content processing.

Operational Environment and Characteristics: the Georgia Archives must establish best practices and standards for content received during the ingest process, including metadata. The system must accept a variety of formats common to state government. The system must transform and validate SIPs according to established Archives best practices and standards.

Ingest accepts SIPs created from transferred, converted or harvested content and transforms the SIPs into Archival Information Packages (AIPs) and Access Content Packages (ACPs) containing metadata and other information required to preserve the objects. These are transferred to content processing. Decisions regarding scope, authentication, versioning and persistent naming are made at ingest.

5.4.1.1 Capabilities, Functions, and Features of the Proposed DAG

- Accepts conforming SIPs
- Rejects non-conforming SIPs and notifies the Archives that content was not ingested
- Makes scope determinations on SIP
- Validates that SIP metadata meets minimum submission level requirements
- Authenticates SIP as official and/or authentic as applicable
- Accepts and transfers integrity marks associated with the SIP
- Accepts and transfers version information associated with the SIP
- Performs version control by detecting if ingested content is already located with the system
- Accepts or assigns a unique ID
- Can accept or assign a persistent name
- Creates metadata

- Transforms SIP into AIP and passes AIP to storage
- Transforms SIP into ACP and passes the ACP to storage

5.4.2 Content Processing/System Administration

Content is processed in a centralized fashion within the system. Rules for authentication, version control, access, and preservation are executed with content processing. Centralizing this processing functionality supports a structured Storage Management System (SMS) and streamlines the required processing.

Content processing accepts, processes, manages, and delivers content. Access Content Packages (ACPs) and Archival Information Packages (AIPs) are key content elements that serve to develop Dissemination Information Packages (DIPs) for content delivery as well as the basis for the package supplied to the Preservation Component for storage.

System administration is the link to the agencies, the end users, and Archives business management functions. It provides the necessary support for processing content to meet the agency and end user needs as well as the business process needs.

As shown in the below figure, the proposed system can be drawn to illustrate the overarching responsibility of System Administration and Content Processing. System Administration ultimately coordinates the ingest of content and its delivery.

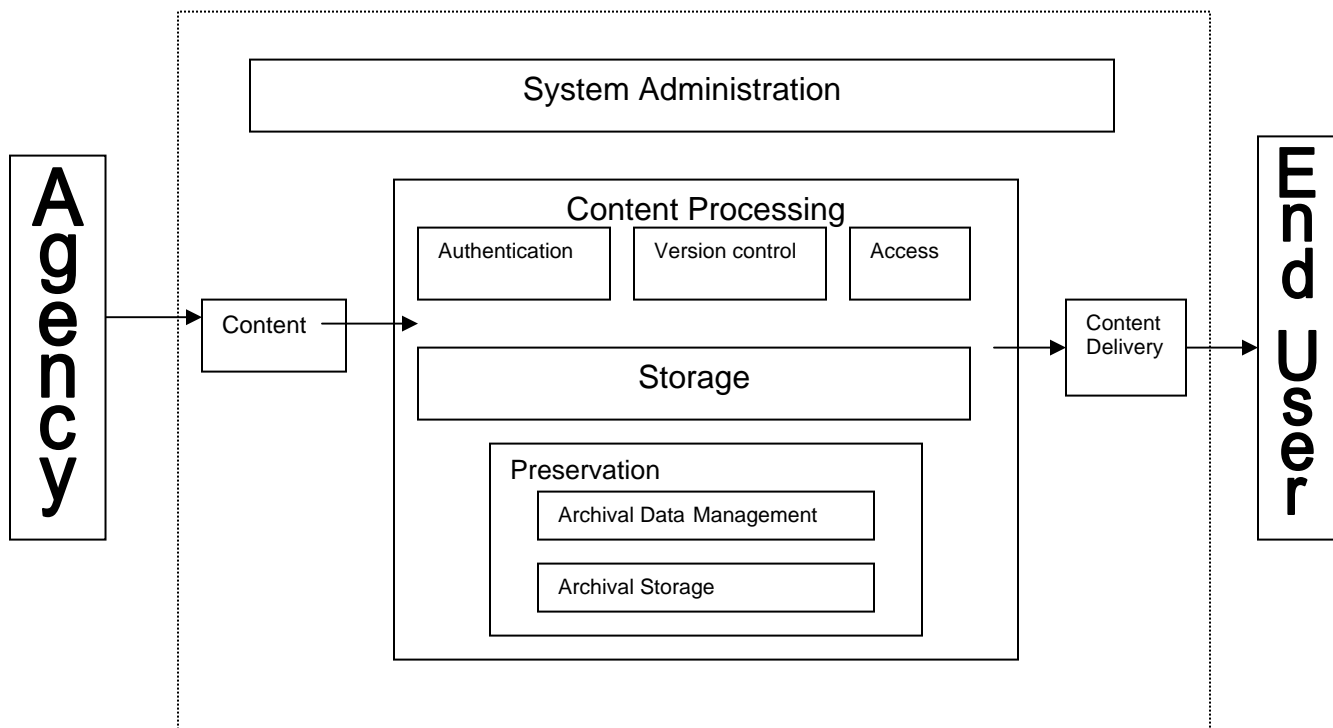


Figure 7 System Administration/Content Processing Component

5.4.2.1 Interfaces to Systems and Procedures

System Administration: System Administration provides the necessary support for processing the content to meet end user need as well as business process needs. In addition, it functions as the workflow controller for the system. In this role, System Administration provides instructions to the other functional components in the system to process content.

Content Processing: The system accepts, processes, manages, and delivers content, AIPs, and the key content elements that serve as the 'master' packages for content delivery as well as the basis for the package that is sent to Preservation for storage.

Content processing includes the rule-processing elements of Authentication, Version Control, and Access as well as the Content Processing and Preservation.

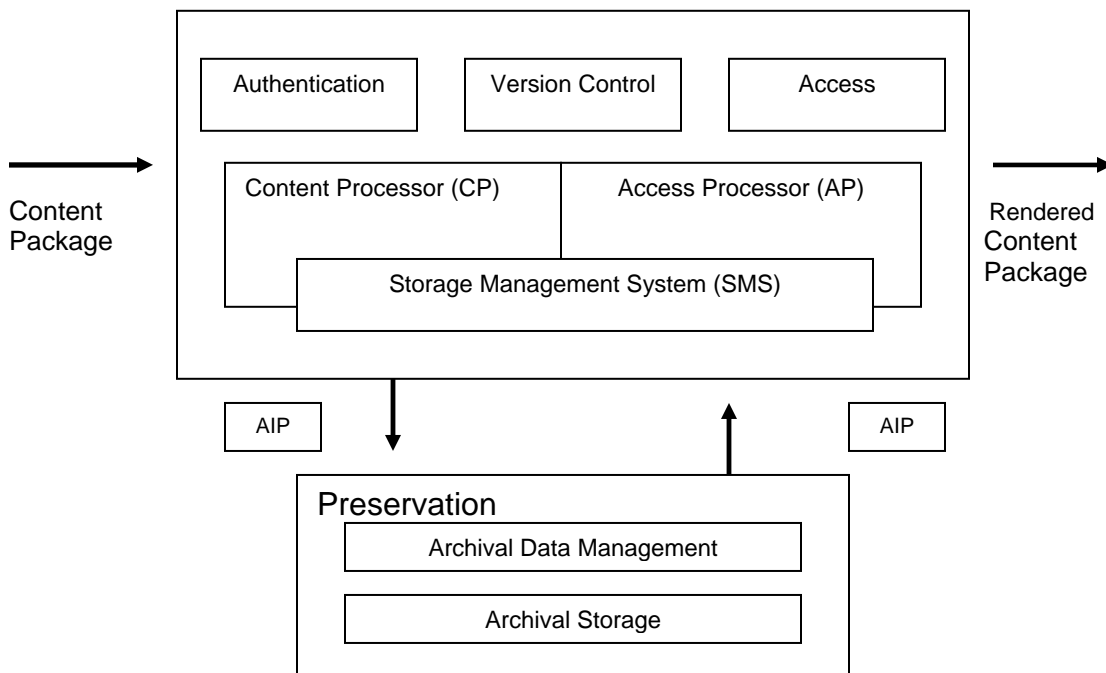


Figure 8 Content Processing

- Authentication – authentication will be called upon to analyze content packages and assign the appropriate attributes.
- Version Control – version control analyzes the content packages and assigns the appropriate version, including version triggers and chain of custody.
- Access – access locates and analyzes the ACPs and develops catalog and index entries for the content packages. Access also delivers the access functions, including search, cataloging and reference tools, request, user support, and interface. Data mining provides access to BPI.
- Preservation – preservation executes the digital preservation processes according to the rules and policies established and maintained within the System Administration component.
- Access Processor (AP)⁷ –
 - Receives requests for Content Delivery via Access
 - Provides request to SMS for retrieval of ACP to be used to create Descriptive Information (DI) in the Content Processor (CP)
- Content Processor (CP)⁸ –
 - Receives and processes content packages

⁷ Refer to figure 6, Proposed DAG metadata model, on page 25, and figure 8, Content Processing, on page 35.

⁸ Refer to figure 6, Proposed DAG metadata model on page 25 and figure 8, Content Processing, on page 35.

- Creates AIP by collecting the Preservation Descriptive Information (PDI)
 - Creates ACPs consistent for Access
 - Creates DIs to meet Access requests
- Storage management System (SMS) –
 - Receives AIP from the CP
 - Receives ACP from CP
 - Manages pre-ingest WIP stores
 - Manages content processing storage (temporary storage)

5.4.2.2 Unique Identifier

Unique identifiers are character strings that uniquely identify all content with the system throughout the content lifecycle. Content managed by the system will be assigned an identifier that exists only once and thus is linked indefinitely to the corresponding content. The uniqueness of the assigned identifier ensures that it refers to only one object.

5.4.2.2.1 Capabilities, Function, and Features of the Proposed DAG

The system will create and assign unique identifiers to:

- Digital objects – upon receipt into the system
- Content packages – including both ACPs, and AIPs
- Job orders – upon request for copies of objects within the system

In addition, the system must have:

- Ability to create and assign an alphanumeric identifier (ANI) for each unique digital object
- Ability to create and assign a unique identifier to a related or continuous piece of content in context
- Ability to record unique identifiers in metadata

5.4.2.3 Persistent Name

In order for the digital content managed by the proposed DAG to be easily found and shared by a wide range of users with different needs and using different systems, there must be a simple way of reliably and unambiguously identifying each resource independent of its location.

Operational Environment and Characteristics: Persistent naming allows for an interoperable schema of identifiers that uniquely identify content, support access to that content, and support access to information about the content.

Major System Components and High Level Interconnection: A resolution system locates and provides access to content and metadata associated with assigned persistent names. The system must be able to associate persistent names to existing legacy naming schema.

Interfaces to Systems and Procedures: The system will assign persistent names to content packages at ingest. All assigned persistent names will be recorded and used in metadata. Once assigned, a persistent name cannot be reused within the system. Procedures will be necessary for updating the location information associated with the identifier if the content has been moved or removed from the system.

Capabilities, Functions and Features of the System:

- Ability to create and assign persistent name at ingest
- Scalability in terms of persistent name assignment and resolvability
- Support for resolution of a single persistent name to multiple distributed locations and content versions
- Capability to log persistent name transactions

5.4.2.4 Authentication

The authentication function within the system will verify that the digital object is authentic and/or official. It will also have the capability to certify the content as authentic or official to others. The term authentic is defined as being complete and unaltered. The term official denotes the 'record of retention,' the copy voted or approved as documentation of an event, action, or decision.

Authentication and certification are an integral part of the data and content management portions of the system. These functions affect all parts of the system, from ingest of a SIP through delivery of an access product via dissemination techniques. This section also addresses granularity issues as they relate to the certification process.

Operational Environment and Characteristics: According to the definition, a document could be considered authentic but not official, but a document could not be considered official and not authentic. The Georgia Archives verifies that the information disseminated through the Archives is official and/or authentic and, when requested, certifies this status to end users.

The system will utilize industry standard technology in order to ensure these relationships. An example would be the implementation of Public Key Infrastructure (PKI). The state of Georgia is currently reviewing authentication standards. A decision on which technology to implement in the DAG will be determined following approval of the standards.

Implementing digital signatures would allow customers to determine that the files are unchanged since they were authenticated by the Archives and help establish a clear chain of custody for electronic documents.

Major System Components and High Level Interconnection:

- The proposed digital archives system must have the capability to verify the authenticity of the content, determine its status as an official document, and certify this to end users.
- The original harvested or converted content ingested will be verified as authentic but not official.
- The system must be able to certify to end users that content is authentic and/or official.
- When encryption standards change, the Archives must have the capability to re-certify content at the higher industry standard encryption levels in order to keep digital signatures (where present) valid and secure.
- The system must provide an automated validating mechanism/process by which the end user would be able to view digital certification information.
- The system must support the digital certification of files in all formats, static and dynamic.

5.4.2.5 Version Control

A version is a unique occurrence of a document or record. The activity of inspecting content for changes and denoting a different version is called version detection. Version information should be reflected in the document metadata.

Version control within the system will be a process of evaluating digital content throughout its entire life cycle. The chain of responsibility will be reflected in the metadata.

Operational Environment and Characteristics: The Archives must determine, either from information provided by the agency or discerned from the content itself, whether the content is a unique occurrence. The Archives must develop policies and procedures to define versions for all content packages, including all types of content. Best practices will be applied at all stages of the system process from ingest to delivery.

Major System Components and High Level Interconnection: Version control will be performed as content is processed.

Content Submission: Ideally, version information is included in the metadata by the agency. This includes reference to other content versions, including hard copy versions.

Content Processing: Apply version control rules to assess the content package and determine content version within the system. Version control tools should be in place to inspect new content packages and monitor existing content packages within the system.

Access: This function will provide information about versions to the user, including version designations, relationships between versions, and chain of responsibility information.

Request and Content Delivery: The selected version must be delivered to the user with appropriate version information attached to indicate that the correct version was delivered.

Administrative Function: The system will allow authorized users to view version information and have responsibility for managing and preserving all versions and application of best practices.

Preservation: The system will preserve all versions of content packages using hierarchal storage management.

Capabilities, Functions and Features of the System:

- Perform version control by detecting if ingested content packages are identical to existing content packages within the system;
- Assign and record version identifiers in metadata
- Apply rules for version control
- Manage all versions of content within the system

5.4.2.6 Access

Access will be the primary interface between end users and the system. Access has been divided into the following functional areas:

- Search – performing queries on content and metadata so that content may ultimately be retrieved from storage and delivered to users
- Request – requesting delivery of content and metadata
- Cataloging – adding metadata to content in the form of standard bibliographic data
- Reference tools – compiling lists that point to content and resources
- Interface – creating user and system interfaces for all functional areas, as needed
- Support – supporting access to content and services

All user classes will be able to interact with the system through Access. While search will primarily provide access to content packages, data mining will provide access to BPI. Access will provide information to users on content package authentication and version control. The system may provide multiple layers of user access and store user preferences to support Access. Access must provide open and interoperable access to content packages.

5.4.2.6.1 Search

Search executes queries on content packages and select external information. The process of searching involves creating a query and refining that query until satisfactory results are achieved. Query refinement may include combining input queries with stored queries and preferences. The enterprise search tools must handle user searches of metadata and content both simultaneously and separately across multiple storage levels and internal and external devices. Search tools should produce a highly relevant, usable, and detailed results list that includes the location and description of meaningful information.

Operational Environment and Characteristics: Search should return the results of the user's query and provide ways to access related information. The Search tools should be designed to meet the needs of all user classes who will be searching the system. Users should have the ability to perform a search on content as well as the metadata that describes that content. Users may require the ability to rank and filter items in a results-set to further meet their needs.

It will be necessary for Search and associated functions to conform to international standards that promote interoperability among networked systems. Search should meet or exceed industry standards for search and retrieval technology.

Major System Components and High Level Interconnection: Search must include accessible and customizable graphical user interfaces that allow all users to submit and refine queries and view and export results sets. Search must conform to international standards for search interoperability.

Capabilities, Functions, and Features of the System:

- Performance: search must perform at or beyond industry standards for search and retrieval technologies
- Standards Compatibility: search must support the archival standards for describing records, MARC, EAD, and other relevant standards.
- Interoperability: Search function should conform to search interoperability standards. For example, ISO 23950 (formerly NISO Z39.50) compatibility provides for interoperable search across locators for information and collections of information. Conforming to interoperability standards will promote interoperability with Internet search engines and federated search utilities, etc.
- Reporting: Search must be able to interact with other elements of the systems, including System Administration and Data Mining function to log all searches for use in reporting usage statistics. Please refer to Data Mining for more information.
- Search Multiple Formats and Levels of Granularity: the system must have the ability to search multiple media, file formats, (e.g. audio formats, video formats, PDF, ASCII, HTML, XML, and other future formats) and levels of

granularity (e.g. by agency, division, program, and specific staff member, or, by record group or series, document, or data element).

5.4.2.6.2 Request

Request will allow end users to request delivery of content and services available from the system. Request will allow users to discover the cost, if any, associated with the delivery of content and services available from the system.

Operational Environment and Characteristics: Request should accommodate all expected user requirements (e.g., request and view content, allow users to discover the cost of printing or requesting delivery of content, choose delivery options, submit payment for delivery). Furthermore, request should allow users to order and submit payment for services delivered by the Archives. Variable fee schedules will be applied based on user class.

Major System Components and High Level Interconnections: Request tools must work in conjunction with Search, Reference Tools, Description, and Retrieval Tools for content delivery. Request will provide the cost of delivering content that has been located by search, request delivery of content, and facilitate acceptance of payment. The Request, with corresponding location and description of content, will be sent to the retrieval tools within Content Delivery, and the retrieval tools will request content from Content Processing. Request must also integrate with other business processes such as inventory control, order tracking, financial, etc.

Capabilities, Functions and Features of the System:

- Interoperable: Request will have the ability to interact with third party systems such as PeopleSoft for financial tracking
- Fee and no-fee based requests: the system must have the capability to process no-fee based and fee based requests
- Security of Customer Transactions: Request must ensure that customer transactions can be conducted in a secure environment
- Request Content and Services: users should have the ability to request both content and services
- Stored User Preferences and Request History: Request must enable customers to store and access user preferences and order histories in a secure environment. The user preferences may contain request status, delivery preferences, preferred payment methods, request tracking, and other customer information.
- Unique Identifier: Request will have the ability to generate a unique identifier for each request.

5.4.2.6.3 Cataloging

These tools create descriptive metadata that conform to accepted standards, and support access and delivery of content.

Cataloging is comprised of the processes involved in describing content to identify or characterize it, providing 'entry points' (terms) peculiar to the information or document by which the information can be retrieved. Indexing is the process of compiling a set of identifiers that characterize a document or other piece of information by analyzing the content of the item and expressing it in terms of a particular system of indexing. In Archives context, cataloging and indexing are the processes that produce Finding Aids.

Major System Components and High Level Interconnection: The cataloging activities take place within Access in the Content Processing layer of the system. Archives staff creates descriptive metadata that the system is capable of expressing as finding aids. Bibliographic metadata includes providing and accessing links to content maintained on various sites management by or completely external to the Georgia Archives, such as the OCLC Digital Archive or the Digital Library of Georgia.

The cataloging process will use applicable metadata elements, including metadata created by the agency.

Interfaces to Systems and Procedures: The cataloging activities take place within Access, in the Content Processing layer. Cataloging will use applicable metadata elements that are acquired with SIPs. Ingest will review SIPs for required metadata elements.

Capabilities, Functions and Features of the System: cataloging will draw upon metadata acquired and stored by the DAG and will transform or organize that data into meaningful structures (e.g. MARC, Dublin Core, etc.).

5.4.2.6.4 Reference Tools

Reference tools present information locators to end users, and add value to the end use experience by assisting in the discovery and access process. Reference tools will contain metadata, references to metadata, and references to content. Reference tools will assist users in locating content packages and select external information.

Major System Components and High Level Interconnection: Reference tools will include lists and resources that assist users in locating and accessing content. Reference tools will have the ability to create, acquire and store metadata, references to metadata, and reference to content.

Capabilities, Functions, and Features of the System:

- Reference tools should be applications or lists that assist users in locating and accessing segments of related content and/or metadata within or outside the system
- The interface to the tools will be published such that external tools will be able to function with the reference tools. For example, the reference tools should be interoperable with other industry standard third party reference tools, such as Internet search engines.
- Reference tools will interact with search tools.
- Reference tools will create additional value for the user experience
- Reference tools may be available in a variety of formats, including electronic and tangible formats
- Reference tools will have the ability to manage and interpret metadata

5.4.2.6.5 Interface

Interface will provide users with direct access to content, metadata, and services available from the system. Interface will meet the needs of all user classes. The system should provide graphical user interfaces and system interfaces where deemed appropriate.

Operational Environment and Characteristics: The Archives will develop interfaces for all internal and external user classes that will allow users to perform authorized functions. Users will have the ability to “opt-in” to acquire the capability to customer default interfaces in order to create an environment better suited to their needs and preferences. A default end user interface will be provided to allow citizens to access historical records without requiring them to submit any form of personal identification (note: we may request a zip code).

Major System Components and High Level Interconnections: Interface will be comprised of graphical user interface and system interfaces. Graphical user interfaces will consist of a workbench or set of user tools related to performing a system function. Changes to a user’s workbench will be saved and made available to “opt-in” users in subsequent sessions.

Interfaces to Systems and Procedures: Interfaces will be made available for all system functions.

Capabilities, Functions, and Features of the System:

- The system should provide an interface for all functional elements
- The system should provide for the creation of system interfaces that promote interoperability among networked systems (e.g. APIs)
- The system should provide for the creation of graphical user interfaces

5.4.2.6.6 User Support

User Support will assist in delivering services to users (e.g. interface user assistance, real-time alert services, and relational databases that track user queries and preferences). These services may or may not be delivered in conjunction with Content Delivery.

Operational Environment and Characteristics: User support enhances the Georgia Archives ability to provide services to users, such as answering questions. User support receives and processes requests to be executed within and outside the system

Major System Components and High-Level Interconnections: Major system components will include relational databases, customer relationship management tools, and other tools that are visible through Access.

Capabilities, Functions, and Features of the System:

- Ability for authorized users to input, store, and manage user preferences and queries
- Support a helpdesk and knowledge bases
- Enable alert services

5.4.2.6.7 Data Mining

Tools and processes for the extraction, analysis, and presentation of content to enhance internal and external business efficiencies.

Operational Environment and Characteristics: The Archives will provide intuitive data mining capabilities including selected access to external data repositories (such as an agency data warehouse) that enable authorized users to run queries against related content from multiple agencies as a long-term strategic planning tool.

Major System Components and High Level Interconnections: The major system components of the data mining function include an interface that allows users to submit and refine queries and view and export results sets.

Interfaces to Systems and Procedures: Data mining is a sub-component of Access within the Content Management Component, with interfaces to the System Administration Component and the Content Processing Component.

User queries will be submitted against the system and external data repositories using various automated data mining tools.

Capabilities, Functions, and Features of the System:

- Ability to extract content in multiple formats from the entire collection

- Ability to normalize data based on administrator defined parameters (e.g. identify missing values or metadata, data formats, types and discrepancies, and anomalies).
- Ability to perform analyses on BPI (e.g. cross tabulations, categorization, clusterization, regression analysis, data patterns and relationships).
- Ability to present data according to user preferences and Archives business rules (e.g. views based on access levels, exporting results, linking of results to data).
- Ability to data mine content within the system at multiple levels of aggregation and granularity.

5.4.2.7 Preservation Services

Preservation services will enable comprehensive, timely, permanent public access to the official records of Georgia government.

Operational Environment and Characteristics: The proposed DAG will manage preservation processes for archived content. Preservation copies of digital objects, AIPS, with associated metadata, will be maintained in Archival Storage.

The Georgia Archives will provide online public access and other services derived from the digital preservation masters in Archival Storage. There is no direct public access or use of content in Archival Storage. Access Content Packages, sometimes referred to as derivatives, of the stored digital objects will be available for online use by the public as well as for print-on-demand and other dissemination methods.

The Georgia Archives digital archives and preservation objectives are achieved by developing and/or following best practices that comply with an adequate, coherent, and widely understood framework of reliable, accountable, and manageable digital archives.

Archival Storage will consist of collections of AIPs with identical content located at several sites. This redundancy ensures guaranteed preservation in the event of a disaster or significant discontinuation of service at a single site.

Major System Components and High Level Interconnections: Archival Storage preserves the content of the historical digital records of Georgia government. The specific preservation processes required by the Georgia Archives are a policy determination. The proposed digital archives system must be capable of supporting activities necessary to keep content accessible and usable, including migration, refreshment, and emulation.

Interfaces to Systems and Procedures: Archival Storage receives AIPs from the ingest function in the Content Processing layer. Preservation processes take

place within the Content Processing layer, and the resulting AIPs are maintained in Archival Storage.

Archival Storage will interface with Content Processing to receive AIPs for preservation and to provide preserved content for the creation of derivative products, including Access Content Packages.

Capabilities, Functions, and Features of the System:

- Content is passed between Content Processing and Archival Storage
- Content is stored in Archival Storage
- Preservation processes are executed in Content Processing

5.4.2.8 Storage Management

This section covers the Access Storage Management for the system.

Operational Environment and Characteristics: Storage management is a key component of Content Management that must provide and coordinate access, backup, and archiving of authentic and official government information as well as ensure data reliability.

Storage management will consist of facilities that are scalable and support increasing and changing storage requirements.

Major System Components and High Level Interconnection:

- Storage system management: this component handles the actual management of the hardware storage devices
- Storage media management: this component handles various media of the system for performance optimization
- Back-up management: this component is responsible for maintaining the system back-up and recovery

Capabilities, Functions and Features of the System:

- Storage must be accessible from several sites
- The Storage Management functionality must support rules/policies for retrieval
- Storage accepts and delivers AIP, ACP and Business Process Storage information (BPI) from the CP
- Storage facilities must ensure uninterrupted availability to content packages
- Executes hierarchical storage management on content packages
 - Near line storage for objects that need fast access
- Manages WIP stores for pre-ingest functions.

5.4.2.9 Security

Operational Environment and Characteristics: The proposed DAG must provide appropriate confidentiality, integrity and availability controls for system information and processes. The system security capabilities are derived from the requirements for authentication, authorization, administration, and auditing. Security requirements include enforcing restriction on access to content, assigning users rights, and maintaining system security. In addition, security requirements include those capabilities that also support the necessary technical controls, operational controls and management controls for the system, such as monitoring capabilities for security. Security concerns both internal and external system interfaces, as well as operational processes associated with the system.

Major System Components and High Level Interconnection:

- User access controls
- User privacy controls
- Confidentiality
- System administration and logging

Interfaces to Systems and Procedures: The security function works across the system. The security function principally interfaces with the System Administration and Workflow to support user access in accordance with policy.

Capabilities, Functions and Features of the System:

- The system shall have the capability to support identification of users
- The system shall have the capability to support user access in accordance with the user's authorization rights and with system security policy
- The system shall have the capability to assure integrity of content within the system
- The system shall have the capability for authorized security administrators to set and maintain system security policy
- The system shall provide the capability to conform with state privacy laws and regulations and Archives privacy policy
- The system shall have the capability to provide confidentiality of user data
- The system shall have the capability to reconstruct complete transactions

5.4.3 Content Delivery

Content Delivery refers to mechanisms for delivering content in a method or manner that fits the requirements of the end user. The delivery methods include hard copy, electronic presentation, and digital media.

- **Hard Copy Output:** all hard copy output from the system will contain a watermark authenticating the copy and providing citation information.
- **Electronic Presentation:** electronic presentations are dynamic and temporary representations of content in Archival Storage.

- Digital Media: digital media consists of the use of data storage devices, such as CDs, DVDs, etc.

Retrieval, a subcomponent of Content Delivery, is defined as the tools and processes that retrieve content from storage and provide transformation information to Content Processing. Retrieval provides a bridge between Request, Storage, and Content Processing. Retrieval retrieves requested Content Packages from storage and passes the Content Package and Request instructions to Content Processing that will be used to transform an ACP into a DIP.

5.5 Modes of Operation

The modes of operation for the proposed system as currently known are:

- Nominal
- Degraded
- Maintenance
 - Remedial maintenance
 - Preventative maintenance
 - Code upgrades
- Alternate site

Nominal mode of operation describes the system when working at the optimum, i.e., the system is operational and working as intended.

Degraded mode describes the system when working at the minimum, i.e., the system has malfunctioned and is working using a reduced string of operations.

Maintenance mode describes operations in time when the system is working using a reduced string of operations. For example, the system is placed in maintenance mode in order to perform a software upgrade. Once the software has been loaded, tested, and verified to work, the system is placed back in the nominal mode.

Alternate site mode can be described as occurring when one site has a failure that requires a user to access records from an alternate site.

5.6 User Classes and Other Involved Personnel

The following subsections describe the classes of users, including user capabilities, which are associated with the system. It should be noted that each user is described functionally by their duties rather than by position title. These descriptions do not represent the number of staff required to operate the system.

5.6.1 Profiles of User Classes

A user can be defined as anyone who will interact with the system. A user class is determined by the ways in which the user interacts with the system. The major user classes are:

- Agency – creator of government records and responsible for their transfer to the DAG.
- Content Evaluator – determines if the records being transferred are eligible for preservation in the DAG by determining if the records are identified as permanent by an approved retention schedule. The content evaluator establishes the initial preservation plan and determines/makes decisions on what processing will occur on the records.
- Content Processor – responsible for ‘processing’ of content including addition of metadata, creation of finding aids, and preparing the content for storage in Archival Storage.
- Systems Administrator/Operations Manager – systems administration directly supports the overall operations and integrity of the system and its use and conducts such system activities as managing user access rights, monitoring system performance, and scheduling reports. The operations manager interfaces with Archives personnel and makes decisions, including approval of workflow processes. The operations manager has ultimate responsibility for the completion of tasks and the quality of the products.
- End User – uses the system to search for and access records, to submit data requests, requests for assistance via mediated searches, and invoke system services.

5.6.2 User Class Capabilities

High-level system capabilities correspond to specific tasks and the users’ needs for the proposed system. These capabilities are organized according to user class, but some capabilities cross class boundaries. **User classes do not correspond to position titles, or single individual users.**

Agency:

- Creates content
- Transfers content and related metadata

Content Evaluator:

- Determines if content eligible for transfer
- Ingests content
- Determines what processing is needed
- Determines the content version
- Develops initial preservation plan

Content Processor:

- Coordinates harvesting and format conversion
- Adds metadata to content (in addition to agency provided metadata)
- Researches context of records creation
- Creates finding aids
- Completes preservation plan
- Moves content into Archival Storage
- Provides technical support for the preservation of content
- Provides end users assistance in searching content (if requested)
- Fulfills end user requests for content in formats other than electronic presentation

Systems Administrator/Operations Manager:

- Establish and maintain user accounts
- Schedules reports as needed
- Uses tools to monitor the system
- Uses the system tools to develop workflows (at the request of staff)
- Modifies workflow when bottlenecks in the system occur
- Quality function administration

End User:

- Uses the system Access tools to locate content
- Uses the system tools to retrieve content
- Uses the system tools to customize content
- Uses the system tools to order content
- Uses the system tools to access selective BPI (e.g. profile information, order status)
- Works with Archives when needed to locate and order content

5.6.3 Interaction among User Classes

The proposed DAG described herein is an overall conceptual workflow model that depicts where user classes should interact within the system and with each other. The figure below illustrates this conceptual model. The Administrator/Operations Manager User and relevant capabilities are embedded in all components of the proposed system.

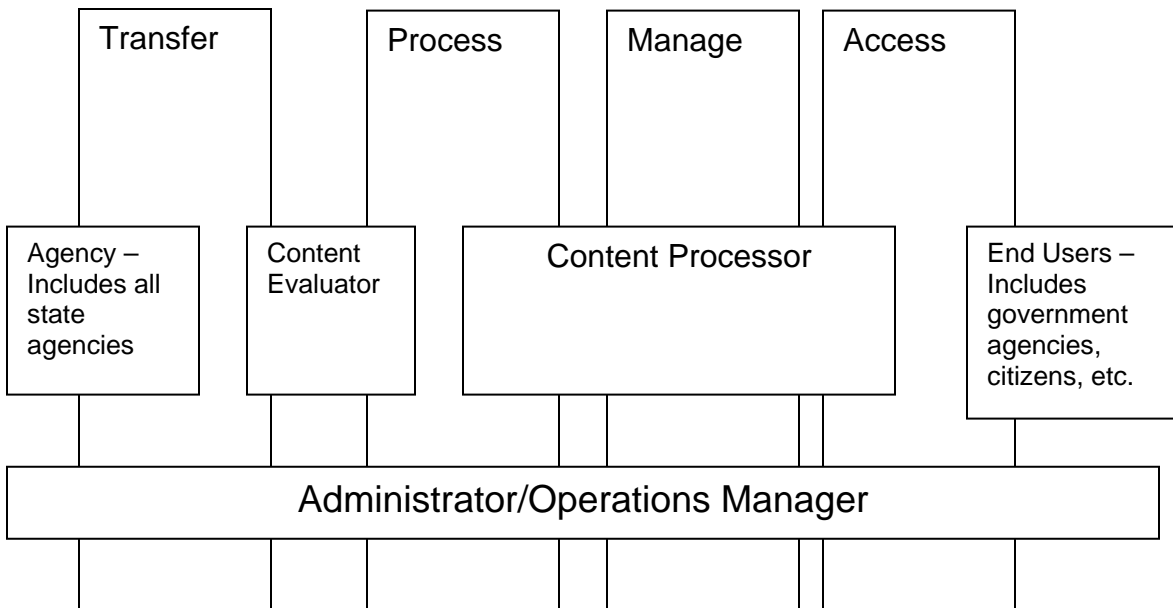


Figure 9 User Classes

5.6.4 Other Involved Personnel

The Georgia Archives will be developing an organizational structure for overseeing the proposed DAG program. When this information becomes available it will be included in a future update of the Concept of Operations document.

5.7 Support Environment

The support environment will not be determined until the conclusion of the systems analysis and design phase of the proposed DAG. However, if another update of the Concept of Operations document is determined to be of value at that time, the required information will be provided.

6 Operational Scenarios

This Concept of Operations document expresses a vision for the proposed DAG. Scenarios convey the system processes in non-technical language. Overlap occurs between different scenarios as a result of interaction between different user classes or due to similarity between different activities. The scenarios below represent one example of how users may interact with the system. They are not intended to identify all possible situations for any given user class. Additionally, the steps in the scenarios should not be interpreted as a fixed sequence of events; instead they serve as an illustration of capabilities the system will offer

(any user class). **User classes do not correspond to position titles, or single individual users.**

6.1 Transferring Agency

Transferring Agency users are creators, custodians, managers and system administrators who create, received, maintain, and manage records. Each agency will have one or more individuals authorized as records custodians or records managers to transfer content to the digital archives.

6.1.1 Scenario -Transfer

The DAG will provide the following levels of support to transferring agencies:

- Accept digital content from the agency regardless of format or characteristics
- Offer guidance and tools to assist in preparing records for transfer
- Allow the agency to access the system in order to search for templates that are needed or to register and store templates
- Manage the workflow process for the transfer of records

The Georgia Archives expects that the DAG will rely on template and related format standards to manage digital content transferred into its custody. In most cases, the level of control and preparation required by the content will vary, depending on the value of the content and the manner in which the content was created and maintained by the agency. In a few cases, the transferring agency will no longer exist, necessitating that the Archives acquire records that it knows little about and which may not meet formatting standards. In this case, Archives staff will act in the role of transferring agency.

6.1.1.1 Activities

1. Develop a memorandum of understanding (MOU) outlining the content, context, structure, and presentation of records and templates to be transferred to the digital archives. The MOU establishes a schedule upon which content will be transferred. Under terms of the MOU the agency agrees to provide all required information and to assist (as needed) in the development of a preservation plan for the content.

- The agency defines templates for sets of content (e.g. case files) as well as for individual types of documents (executive orders, legislative bills) and registers them with the system. The agency accesses the template repository to utilize common templates, register a new template; or to modify an existing template. Agencies may search the template repository to determine if a template exists for a particular document. The system will provide a number of common templates for use in transferring various

types of records. These common templates will contain only the minimum required descriptive elements and would be used transfer the records of a defunct agency (Archives staff handles) or to transfer older digital content where it is not appropriate to develop a unique template.

- The DAG checks the templates to ensure they meet Archives standards and include all mandatory elements. If the template is rejected, the system will notify the agency of the problems and request correction.
- The Agency selects a method for transferring the content and identifies any specific requirements it has for retrieving the content after transfer.

2. Transfer Content

- The Agency will initiate a request to transfer content. The DAG will associate the transfer with registered templates. If the template matches the transfer, the transfer will be approved by the system. Upon approval the content is transferred. For transfers that are denied because they lack information or are incorrect in some manner, the system will notify the agency of the problems or request that the agency contact an Archives staff member for assistance.
- The Agency will send the content, along with the required supporting information, to the system using a reliable method of communication. The DAG will provide the capability to accept content transferred via telecommunications or on acceptable digital media.

Note: the DAG is not intended to manage the transfer of paper or microform materials.

6.2 Content Evaluator

The content evaluator works with agencies to develop memoranda of understanding establishing a schedule for content transfer. In addition, the evaluator reviews content to ensure that it is eligible for transfer to the digital archives (covered by an approved retention schedule identifying the content as permanent). The evaluator is also responsible for gathering information for the preservation plan.

6.2.1 Scenario – Transfer and Ingest

The DAG will provide the following levels of support to transferring agencies:

- Ensure that content is associated with a registered or common template
- Streamline the review and approval process for transfer
- Manage the workflow process for the transfer of content

In cases where the transferring agency no longer exists, the content evaluator will act in the role of the transferring agency in addition to performing the role of content evaluator.

6.2.1.1 Activities

1. Review memoranda of understanding and related templates, and produce the Preservation Plan.

- The Agency requests the transfer of digital content based on the terms of an MOU. Using the tools in the system, appropriate templates are selected and required information is provided in support of the transfer. The system performs an initial review of the provided information to ensure that it meets Archives standards.
- The system checks the templates identified for the specific content to be transferred. If the transfer is rejected (based on the templates or related information provided), the system notifies the Agency that the documentation is in error. Once the transfer is 'accepted' by the system, a content package comprised of the MOU, any related information, templates, and their respective validation reports is created by the system and the evaluator is notified of a transfer pending review.
- The evaluator accesses the system to review the content package. The evaluator determines whether the information provided is adequate to accept the content. Based on the type of content that is proposed for transfer the evaluator determines if there are standard methods which would enable transfer, preservation, and access to the content. If so, these are specified in the preservation plan. In any instance where there is no standard method, the evaluator contacts a content processor to identify an appropriate method.

2. Review and evaluate records and related information.

- After the evaluator determines that the content package is adequate (compliant with applicable policy and procedures), the evaluator may contact the Agency in order to examine the content to be transferred. The evaluator may either request that a sample of the content be transferred to the system or that the evaluator may have access to review the content in the agency systems. In some cases, review of the content may not be required.

3. Notify Agency of transfer status.

- In the event that corrections to the content package are required, the system will maintain an event log to track the status of the transfer approval process.

- The system will enable the evaluator to notify the Agency of the status of the content package.
- The system provides authorization to transfer records based on the successful evaluation of the content package.

6.3 Content Processor

The content processor administers the accession, verification, arrangement, and description of digital content. The processor may interact with the Agency and with other Archives staff roles at various points in order to gather information for the creation of finding aids. Content processors may have special technical skills that enable them to focus on the preservation activities, while other processors will focus on the arrangement of content and the creation of metadata and finding aids.

Preservation activities include those activities necessary to ensure that digital content remains intact in transfer and storage; that a digital object can be reproduced from its components and presented in authentic form; that the original order, and any other arrangement of digital objects can be implemented; and that digital content may be certified as both authentic and official.

6.3.1 Scenario - Process

The DAG will provide the following levels of support to the Georgia Archives:

- Provide for verification of content upon transfer
- Provide tools to streamline the processing of content
- Provide tools to enable the preservation of content
- Facilitate the creation of finding aids through the use of templates
- Manage the workflow process for the creation of finding aids

6.3.1.1 Activities

1. Transfer

- The Agency transfers the content.
- The system stores the content in a manner consistent with its stated level of security. The system confirms receipt of the transfer and notifies the content processor of the arrival of content.
- The processor identifies characteristics of content that cannot be accessioned, preserved, or accessed using existing tools or templates, and determines if they could be accommodated by modification or extension to existing tools and templates, or by creation of new templates.
- The content processor uses the system to check that transfers are successful (i.e., checks that all the appropriate components of the transfer are included and that they are complete and uncorrupted). The system

- produces reports of any problems or discrepancies and notifies the processor and the Agency. When the Agency has transferred any missing, incomplete, or corrupted components, or has replaced such items, the system notifies the processor.
- The processor articulates the preservation standards that guide the evaluators in working with the agency to transfer content.

2. Verification and Security Review of Content

- The processor invokes a preliminary screening of the content to identify or check for the presence of potentially restricted content. The system scans the content for indications of restricted information and reports the results to the processor. The system will segregate restricted content for further review.
- The system checks the content against the retention schedule, transfer authorizations, preservation plan, and other documentation. The system produces a report showing the results of the verification and includes the report in the accession package.
- If the verification report indicates a discrepancy, the processor is alerted. The processor determines whether to add or note in the documentation or in the description and describes the discrepancy. If necessary, the processor may also notify the Agency and request corrective action.
- The processor indicates the successful completion of the verification process.

3. Preservation

- The processor works with administrative users to ensure that the system can implement standard preservation methods and check transfers against these terms and conditions.
- The system checks that the methods detailed in the preservation plan can be applied successfully to the content.
- If the preservation plan requires any actions, such as reformatting data types or record types (i.e., transformations), the system stores the content as received and performs the action, generating a second version of the content.
- The system carries out the instructions of the preservation plan and generates a report on preservation verification and actions, identifying any problem encountered.
- The processor reviews the report and may decide to run tests to check the authenticity of the content.
- The processor defines requirements for an audit trail of all transformation performed in the system in order to document the relationship between the content acquired from the Agency and their transformed versions, and defines reporting requirements for other system functions and parameters related to preservation.

4. Arrangement

- The processor uses the system to determine the arrangement of digital objects in the transfer and to check that the system can substantiate this arrangement. The system must be able to maintain the specific arrangement of the objects and must be able to present the objects in the specified arrangement. The system will store the specification for an arrangement in a template. Digital content may be arranged:
 - In a hierarchical files classification
 - According to a data model
 - Empirically on a web site
 - In layers of a geographical information system
- The processor determines whether a template specifying the original order of the records already exists in the system. If so, the processor tests to confirm that the system can present the objects in the arrangement specified. If not, or if the template is inappropriate, the processor may define a new arrangement for the objects. The system transmits any proposed new arrangement to the appropriate manager for approval. Upon approval, the template is stored as an approved arrangement template.
- The system will provide the capability to support multiple arrangements and manage versions of arrangements. Hierarchical levels of arrangement may be specified in a single template. The system must also be able to link templates, so that, for example, a single template applies to a series which exists over long periods of time, but parallel, lower level templates reflect successive changes in the internal structure of the series.
- The processor may use the system to define relationships between or among objects or content sets, in addition to the relationships stipulated in an arrangement. For example, the processor may indicate that the same document is a document in two or more different file units or series. This occurs when many records in a case file of an agency are duplicated in the files of subsequent investigations.

5. Accessioning (Transfer of Legal Custody)

- The system assembles the information relating to the content to be accessioned, including transfer documentation, the description of the content transferred, correspondence, and any other relevant information. A processor uses the system to complete and submit the appropriate document transferring legal custody.
- The system notifies a manager that the accession is ready for approval. Designated staff review the accession information, the content (if needed), and approve the transfer of legal custody. When approved, the system indicates that the accession has been accepted and notifies the Agency.

6. Description

- The processor uses the system to create or enhance the description of content. The system parses the transfer documentation, the digital content, and reports generated in processing the records, extracting pertinent information that will populate description fields. Any required fields not populated will be completed by the processor. The processor reviews and modifies the description as needed. A manual description options exists that enables the processor to enter information manually or to copy and modify existing description. When the processor indicates that the description is complete, the system checks the draft against business rules for archival description and notifies the processor of any problems. The system will manage descriptions and provide maintenance of multiple versions of descriptions.
- The processor submits the completed description for review to a manager. The system manages the approval workflow.

7. Conclusion of Processing

- The system notifies the processor that the accession is complete.
- Content is stored in Archival Storage.
- The system notifies end users that content is available for access.

8. Maintenance of Digital Content

- The processor uses the system to examine samples of objects being preserved to ensure that nothing is lost or corrupted in storage. The system will provide the capability to monitor raw bit error rate and corrected bit error rate of storage media. The system will also provide safeguards to monitor media degradation, migrate to new media, and provide the necessary tools for recovery of objects from failed media.
- The processor works with administrative users to ensure that necessary changes, such as media migration, are implemented in the storage system. The processor reviews plans for, monitors, and evaluates updates or modifications of the storage system, including migration of preserved objects to new media.
- The processor identifies opportunities for improving preservation quality or services and uses the system to perform such changes.

6.4 Administrator/Operations Manager

Administrative users are those that handle such activities as assigning user rights and privileges, scheduling reports, monitoring the system, modifying workflow, and ensuring system availability.

Operations managers are those users who are responsible for making decisions related to the lifecycle management and processing activities and assignment of personnel to perform the work scheduled. Responsibility for completing the tasks rests with the operations managers. The manager also interfaces with the administrator when system problems disrupt the flow of work.

6.4.1 Scenario - Administrator

This scenario is included to demonstrate some of the capabilities that would be included in the DAG for administrative users. Not all administrative capabilities are described in the scenario and many of the system functions will be done without user involvement.

6.4.1.1 Activities

1. Assign user rights and privileges

- Using predefined roles (defined by the Georgia Archives and including information regarding clearances, permissions, and job roles), the administrator creates the user account establishing requested access rights and privileges in the system. For example, users with appropriate access will be able to view confidential information necessary for their work. The user is granted appropriate access rights (e.g. access to restricted data or administrative access) and systems capabilities (e.g., ability to edit, input data, check security, produce user reports). Note that users with 'public' access rights can be created by the system, and that user accounts can also be established for those users who wish to avail themselves of fee services.

2. Schedule reports

- The administrator logs onto the system and uses any data available in the system to create new reports or modify existing reports. The request for reports could be based on a specific requirement from or a system monitoring need. The reports could provide metric data for such activities as system usage, system capacity, performance, or workflow statistics. The system provides the ability to manage report (i.e., create, modify, save, delete) and has the ability to output the reports via a user interface, media, or to external systems. The system can also make reports available to other users of the system.
- The reports are scheduled for regular distribution to the appropriate people or are created on as needed basis.
- Schedules and contents of reports can later be changed as required.

3. Monitor System

- The system provides the administrator with the ability to monitor system performance and security using system tools such as a dashboard. The dashboard is an integrated set of diagnostic tools that is used for monitoring the health of the system. It monitors storage, performance, space, load, security-related indicators, etc. When a system status alarm occurs (as indicated on the dashboard) indicating a system problem/fault or a potential security problem, the system alarm log is updated with alarm type/number, and time and date stamp of the occurrence for audit trail purposes. Additionally, the system notifies the administrator of the problem via automatic paging, telephone call, or some other method.
- The administrator diagnoses and troubleshoots problems implementing intrusion detection system and virus control procedures. In parallel, the system is recording these events in system logs and establishing an audit trail.
- Once the problem has been corrected the administrator ensures that the system's operations are secure from intrusion, viruses, unauthorized access, etc., i.e., the system performance is as intended.

4. Modify Workflow

- In some instances, the administrator will be able to modify workflow. This does not mean that the administrator will be able to modify rules, assignments, etc. for the lifecycle workflow using the system. The administrator will be able to modify work flowing through the system at a point in time when problems with the system arise.
- When the administrator is alerted to a potential problem with the system (e.g. a problem with the server has occurred) or has been notified of a problem and workaround recommendation by a manager, the administrator diagnoses and troubleshoots the problem and temporarily modifies system workflow(s) to ensure continued service.
- The administrator notifies the manager of the temporary modification to workflow. The administrator tracks the resolution of the problem for audit trail purposes and the modified system workflow(s) will exist in the system until the problem is corrected.

6.4.2 Scenario - Operations Manager

This scenario is included to demonstrate the interaction of the manager with the workflow capabilities of the system, and the interaction with the administrator. Note that some steps are performed by the system without the need for human intervention, and some are a combination of system and human activities.

6.4.2.1 Activities

1. Job Pending

- The manager logs onto the system and receives a notification from the system that a job is ready for processing. The notification indicates that content is being transferred into the system. The system, using predefined rules, determines what activities need to occur. Based on these rules, the system can decide to create a job, assign jobs, assign due dates, note access restrictions, and provide relevant information about the objects in the job.

2. Review System Assignments

- The manager reviews the assignments identified by the system and selects from the options that are presented:
 - Confirm assignments, or
 - Modify assignments.
- Confirm Assignments
 - Upon confirmation by the manager, the system notifies staff of assignments including milestones and begins to track the job, which include capturing performance statistics
 - As the job proceeds, the system is able to send notifications, collect approvals, detect when processing has been suspended, make additional assignments, or notify the manager that a job is complete.
- Modify Assignments
 - Upon inspection of the job, the manager has the capability to modify the steps, adding or removing steps, or changing the order of the steps to be performed to process the job.
 - The system will either confirm the modifications or may determine that additional steps are required
 - Upon approval, notifications are sent to staff alerting them of assignments.

3. Approval and Closure of Processing

- As the job progresses through the system, there are various points where manager approval is required. The manager will inspect jobs on a periodic basis and provide approval as appropriate, including final approval that the job has been successfully completed.
- If the manager disapproves a job pending completion approval, notification will be sent to staff providing the determination, justification for the decision, and possible remedies.
- Staff will modify the proposed processing and the approved job will be processed per normal course of operation.

- Upon final approval, the system captures this information and stops tracking the job.

4. Modify Workflow

- When notified by the system that identified steps are not occurring as scheduled, the manager has the capability to examine the system in an attempt to understand and/or determine the nature of the problem. Possible problems could be related to bottlenecks in the system or the inability of staff to complete tasks.
- The manager may have to interface with the administrator and recommend possible solutions or interface with staff to determine the nature of the problem and recommend solutions.
- The administrator has the capability to implement an agreed upon solution in order for processing to continue.

6.5 End User

An end user is any individual or organization who wishes to identify and/or obtain access to or copies of digital content that is preserved in the digital archives system. These individuals fall into three broad types:

- Agency staff, including agency content creators, records officers, managers, courts, legislative staff, and others. An agency has unrestricted access to its own records (but may not alter them) and possibly to the records of agencies with which it has a data sharing agreement.
- Archives staff members, including those staff members that perform access review, arrangement and description, reference services, system operations, and others.
- Public, including researchers, historians, genealogists, etc.

6.5.1 Scenario

This scenario describes how the end user will employ the digital archives system to search for, access, and retrieve content. The system will support search of descriptions of all content, and show defined relationships between the sets of digital content and non-digital records; however, those capabilities are not described in this scenario. The system will allow a broad array of search and retrieval capabilities that can be adapted to each user's needs, privileges, and access.

6.5.1.1 Activities

The steps below should not be necessarily interpreted as a sequence of events.

1. Access

- All end users will be able to search and retrieve descriptions of all content accessioned by the Georgia Archives. In addition, they will be able to search and retrieve digital content which have no access restrictions that are maintained in the system. End users with special access rights and privileges may search and retrieve digital content containing restricted information.

2. Search

- End users search the system for information describing content and for actual digital content. This may be done at a variety of levels of aggregation (i.e., record group or set, series, file unit, item, data element). Within the end user's rights and privileges, the user may take advantage of available functions and features. The system responds to queries against descriptions by supplying the descriptions that match the search criteria. Normally, objects are described at the set level, such as series or file unit. If descriptions of interest are identified, the user may proceed to run queries against the content of those descriptions. The system responds to such queries by identifying either sets of objects or individual objects, with results constrained by the user's access rights. The system provides the capability for the end user to view and/or sort the results of the search, modify the search if necessary, and refine or save results as desired. The user is able to perform these functions in an iterative fashion, thus permitting the user to progress from a search about a general topic to a list of specific objects that the user may wish to view.

3. Retrieve/Receive

- From search results that identify relevant content, the system allows the end user to view and access the content desired. The user directly interacts with the system and accesses the content in accordance with established privileges and access rights. The end user may request the system to output content to a selected medium or print them in formats with parameters chosen from available options. The system also provides the capability to direct output via telecommunications, for example, using File Transfer Protocol (FTP). The end user may use search and retrieval capabilities without any involvement from Archives staff, but if at any time the user has questions, has trouble searching, requires services, or is unable to retrieve/receive records due to access restrictions, the system provides the end user the capability to request a mediated search.

4. Mediated Search Request

- End users may request help from Georgia Archives staff while using the system. A mediated request may include such activities as staff answering questions, conducting and handling searches, providing certified copies, processing special requests, expediting requests, handling Open Records Act requests, and similar issues. The system tracks the communication and information about the mediated request. After all questions are answered, issues resolved, and special requests processed, the end user retrieves/receives the objects as described above. If the objects are restricted the end user may instead receive information concerning the status of a particular request.

5. Fee for Service

- End users may request products that require them to pay a fee. If a fee must be collected or charged-back for any special product during this process, the system tracks, reports upon, and routes any required financial transaction information to all appropriate accounting systems, and provides the requested product on authorization by the accounting system.

7 Summary of Impacts

Implementation of the proposed DAG may have wide ranging impacts on both the Georgia Archives and its customers. The subsections below identify potential operation impacts, organizational impacts, and impacts during development that should be considered as the Archives develops its plans.

7.1 Operational Impacts

Until the proposed DAG undergoes systems analysis and design, operational impacts of the proposed system are not known; therefore, impacts to the following have been omitted.

- Interfaces with primary or alternate computer centers
- Changes in procedures
- Use of new data sources
- Changes in quantity, type, and timing of data to be input into the system
- Changes in data retention requirements
- New modes of operation based on emergency, disaster, or accident conditions
- New methods for providing input data if the required data are not readily available
- Changes in operational budget
- Changes in operational risks

However, it is anticipated that the Georgia Archives as an organization will have to implement changes to the way it conducts business in order to achieve the agency's mission, goals, and objectives. The system will facilitate this endeavor. When implemented, the system will be able to handle vastly different content in a variety of formats that the Georgia Archives has been incapable of addressing in the past.

The system should provide decision support for the Georgia Archives management processes for the content management lifecycle processing of all content. This additional capacity includes supporting processes such as evaluating, creating descriptions of and metadata for, and preserving content of all types. Additional operational impacts may include the following items:

- Data architecture modeling;
- Disaster or catastrophic recovery;
- Advances in technology; and,
- Changes to operational procedures.

7.2 Organizational Impacts

The Georgia Archives is examining current policies and business practices and may have to develop and/or modify these policies and practices as system analysis and design progresses. The depth and breadth of the organizational impact is unknown at this time. Information with respect to the following has not been provided for this reason and includes such items as the number of personnel, skill levels, position identifiers, and locations. Additionally, the interaction of personnel with the system may necessitate revising position descriptions to reflect the anticipated changes in the Archives' business practices.

With this in mind, the Georgia Archives has identified a number of possible organizational impacts, as described below:

- An assessment of how the system will fit organizationally within the Archives and how the system will relate to other Archives components and/or the agencies they interact with;
- The commitment of resources (e.g., funding, time, staff) by the Archives to establish working relationships with other government agencies as the system moves from pilot into production;
- The need for cross-functional, inter-disciplinary staff teams;
- The development of education and increased training for both Archives staff and end users;
- The possible need for a help desk facility for Archives staff and end users (such a function would require adding staff); and,
- Improved opportunities for career development for Archives staff.

7.3 Impacts during Development

The full extent of impacts during development will not be known until completion of the systems analysis and design phase, and, as such, this information has not been provided; however, impacts considered thus far include:

- Articulation of business rules and other controls needed for operational implementation;
- Development of training for requirements to be implemented in increments; and
- Training necessary for rollout of each increment.

When known, information on impacts such as the following will be provided:

- Involvement in studies, meetings, and discussions prior to award of any contracts;
- User support and involvement in reviews and demonstrations, evaluation of initial operating capabilities and evolving versions of the system, development or modification of databases, and required training;
- Parallel operations of the new system and processes with existing processes; and,
- Operational impacts during system testing.

8 References

8.1 Standards and Guidelines

The large-scale archival storage of digital objects, Jim Linden, et al, The British Library, Digital Preservation Coalition Technology Watch Series Report 04-03, February 2005, <http://dpconline.org/docs/dpctw04-03.pdf>

Cornell University Library. Digital Preservation Strategies. 2003. Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems. www.library.cornell.edu/iris/tutorial/dpm/terminology/strategies

Cornwell Consultants in Management and IT. Model Requirements for the Management of Electronic Records (MoReq). 2001. Electronic Document and Records Management (EDRM). www.cornwell.co.uk/edrm/moreq.asp .

Council on Library and Information Resources, Report 92, "Authenticity in a Digital Environment," May 2000, <http://www.clir.org/pubs/reports/strategies.html>

Digital Curation Centre (DCC), Digital Curation Manual, <http://www.dcc.ac.uk/resource/curation-manual/>.

Digital Preservation Coalition, Handbook on Digital Preservation,

<http://www.dpconline.org/graphics/digpres/index.html>.

International Council on Archives, Electronic Records: A Workbook for Archivists (ICA Study 16), 2005, www.ica.org.

IEEE 1362-1998: *IEEE Guide for Information Technology – System Definition – Concept of Operations (Con Ops) Document*.

ISO 14721:2003: *Space data and information transfer systems – Open archival information system – Reference model*.

National Library of Australia, Preserving Access to Digital Information (PADI) Initiative, "Digital Preservation Strategies," <http://www.nla.gov.au/padi/topics/18.html>.

Preservation Metadata, Brian Lavoie and Richard Gartner, The British Library, Digital Preservation Coalition Technology Watch Report Series Report 05-01, September 2005, <http://www.dpconline.org/docs/reports/dpctw05-01.pdf>.

Preservation Metadata Implementation Strategies (PREMIS) Working Group. Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group. May 2005. www.oclc.org/research/projects/pmwg/premis-final.pdf

Preservation Metadata Implementation Strategies (PREMIS) Working Group. Official Web Site. www.loc.gov/standards/premis.

The National Archives, United Kingdom, "Sustainable Electronic Records, Strategies for Maintenance and Preservation of Electronic Records and Documents in the Transition to 2004," Version 1.0, August 2001, http://www.nationalarchives.gov.uk/electronicrecords/advice/pdf/preservation_toolkit.pdf

8.2 Laws and Regulations

"Georgia Records Act" O.C.G.A. §50-18-90 et seq

"Georgia Open Records Act" O.C.G.A. §50-18-70 et seq

Georgia Archives enabling legislation O.C.G.A. §45-13-40 et seq